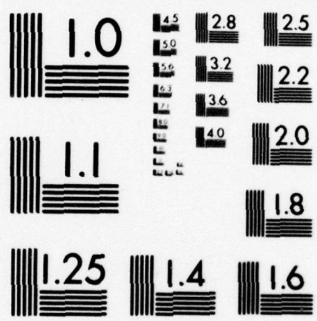
AD-A064 658   AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC   F/G 12/1
CRITERION FOR SELECTION OF VARIABLES IN A REGRESSION ANALYSIS.(U)
DEC 78   L J PULCHER

UNCLASSIFIED   AFIT/GOR/MA/78D-5   NL

1 OF 2
AD
A064658



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL III

1

ADA064658

DDC FILE COPY

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson Air Force Base, Ohio

79 02 02 047

AFIT/GOR/MA/78D-5

1

LEVEL II

6 CRITERION FOR SELECTION OF VARIABLES IN A REGRESSION ANALYSIS.

9 Master's THESIS,

14 AFIT/GOR/MA/78D-5

10 Larry J. Pulcher
Captain USAF

11 Dec 78

12 137p.

012 225 —

bpg

Acknowledgments

This study was conducted at the Air Force Institute of Technology under the guidance of Dr. David R. Barr. I would like to thank Dr. Barr for his patient guidance throughout the process. I would also like to express appreciation to my wife, Lisa, for her assistance and encouragement and to my son, Jacob, for his patience.

Larry J. Pulcher

Table of Contents

List of Figures

List of Tables

Table		Page

Abstract

The Test of Equality between Subsets of Coefficients in Two Regressions is developed and applied as a means to pre-screen variables from a regression model.

Some criterion for selection of variables are discussed and some existing regression packages are applied to data on characteristics of avionics equipment for comparison purposes.

CRITERION FOR SELECTION OF VARIABLES IN A REGRESSION ANALYSIS

I. Introduction

Previous Results

At the request of the Systems Evaluation Branch of the Air Force Avionics Laboratory at Wright-Patterson AFB, the Westinghouse Electric Corporation performed a regression analysis of characteristics of Line Replaceable Avionics Units (LRU) in an attempt to model some of the equipment's logistics characteristics. Westinghouse identified 21 independent and six dependent variables. Seven of the 21 independent variables are purely indicators, identifying type of aircraft and general usage category of the equipment.

The regression was performed using the "Linear Least Squares Curve Fitting Program" (LLSCFP) developed by Daniel and Wood. In using LLSCFP, all 21 independent variables, as well as some terms containing the squares or the natural logarithms of variables are included simultaneously in the first regression. Then with the aid of statistics, plots, and tabular data arrangement, a subset collection of independent variables is chosen which best approximates the data.

Scope of Present Study

There are several ways to arrive at the "best" subset of independent variables to include in a model. At one extreme, the model can consist of all possible variables. But this is not desirable for two reasons. First, it may be very expensive to gather and maintain a data base for a large number of variables, some of which may have little impact. But second and more important, when the purpose is to predict future costs, as it is in this study, a model that uses a large number of variables to fit the nuances of previous data may in fact have a higher prediction variance than a subset model (Ref 17:7). Important information could be lost in the myriad interrelationships that exist. For this reason the selection of the form and the variables of the regression equation becomes important.

The methods of selection of variables to be investigated in this study are: iterative techniques using the Statistical Package for the Social Sciences (SPSS) and BMD Biomedical Computer Programs (BMD), all possible regressions using the Leaps and Bounds technique developed by Furnival and Wilson, and the C_p statistic search using LLSCFP developed by Daniel and Wood.

The desired final outcome of this analysis is to provide the personnel at the Air Force Avionics Laboratory with a method of performing a quality regression without an extensive background in the technique, so that they can do work in-house which they previously

contracted out. For that reason the methods used in this work will rely on existing packages where possible.

Theory of Linear Least Squares Regression

Assumptions. The first assumption made in the use of linear least squares regression is that the correct model has been chosen. If an incorrect form is used some values given by the equation will be biased.

The second assumption is that the data is typical of the true population about which the analysis is being performed.

The third assumption of the method is that the y observations are statistically uncorrelated and independent. If each y value is considered to be composed of a true and a random error value called ϵ, then this assumption can be restated as: The expected value of the product of any two of the random components is zero.

Three other assumptions that are considered less important (Ref 8:8) are that all observations on y have the same unknown variance, that the levels of the independent variables are non-stochastic, and that the uncontrolled error is distributed normally.

Method of Least Squares. The general form of the linear least squares regression equation is

$$y = \alpha_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \epsilon \qquad (1)$$

where y is the observed value of the dependent variable, x_i is the

observed value of the i^{th} independent variable, and α_o and β_i are regression coefficients. While the linear squares method can treat only equations in the form of equation (1), there are non-linear equations which are intrinsically linear, such as $y = \alpha_o x_1^{\beta_1} x_2^{\beta_2}$. By taking the natural log this equation becomes

$$\ln y = \ln \alpha_o + \beta_1 \ln x_1 + \beta_2 \ln x_2$$

This is only one of a variety of intrinsically linear equations and methods for linearizing. Equations that are not intrinsically linear can not be handled using the linear least squares method. Non-linear least squares packages are available but will not be considered in this study.

If there are N independent observations of y_i and x_i, each observation can be written as

$$y_i = \alpha_o + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_K x_{iK} + \epsilon_i \tag{2}$$

where x_{ij} represents the i^{th} observation of the j^{th} variable. If the matrices

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} \tag{3}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{N1} & x_{N2} & \cdot & \cdot & \cdot & x_{Nk} \end{bmatrix} \tag{4}$$

$$\underline{\beta} = \begin{bmatrix} \alpha_o \\ \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix} \tag{5}$$

$$\underline{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_N \end{bmatrix} \tag{6}$$

are defined, then all N equations can be written simultaneously as

$$\underline{y} = X \underline{\beta} + \underline{\epsilon} \tag{7}$$

The basis of the least squares method is that a straight line is generated through data points in such a way that the sum of the squared distances or errors between the line and the points is minimized. This sum of squared errors (SSE) can be written as $\sum (\epsilon_i^2)$ and in matrix notation is

$$\sum (\epsilon_i)^2 = \underline{\epsilon}' \, \underline{\epsilon} \tag{8}$$

Substituting (7) into (8) yields

$$\underline{\epsilon}' \, \underline{\epsilon} = (\underline{y} - X \, \underline{\beta})' \, (\underline{y} - X \, \underline{\beta}) \tag{9}$$

which we would like to minimize. If classical optimization is performed on (9), an estimator of $\underline{\beta}$

$$\underline{b} = (X' X)^{-1} X' \underline{y} \tag{10}$$

is found to minimize SSE.

The variation of the dependent variable measurements about their mean

$$(y_i - \bar{y})^2 \tag{11}$$

is called Total Sums of Squares (TSS). TSS can be decomposed to two components, the Sum of Squares Explained by the Regression (SSR) and the error unexplained by the regression (SSE) such that

$$SST = SSR + SSE \tag{12}$$

This partitioning gives rise to a measure of goodness of fit R^2_{yx} or commonly written as just R^2 and called the Coefficient of Determination or the Multiple Correlation Coefficient Squared. R^2 is defined as

$$R^2 = \frac{SSR}{SST} \tag{13}$$

and represents the fraction of the variability in the independent

variable explained by the regression equation. But it is known that the Multiple Correlation Coefficient calculated in this way is biased upward, always indicating a higher degree of correlation than actually exists in the true population. In order to correct for the bias, R^2 is adjusted for degrees of freedom by the relationship

$$\bar{R}^2 = 1 - (1 - R^2)\left(\frac{N-1}{N-K-1}\right) \tag{14}$$

and called the Adjusted Multiple Correlation Coefficient. While $\bar{R}^2$ is not entirely unbiased, it does exhibit less bias than R^2, and will be used in this analysis.

A third measure of goodness of fit is the C_p statistic derived by Mallows, and is based on total squared error. The total squared error can be considered to be made up of a squared bias plus a squared random error in y at each data point. If the total squared error is represented as

$$\sum_{i=1}^{n} (\nu_i - \eta_i)^2 + \sum_{i=1}^{n} \mathrm{Var}\,(y_{ip}) \tag{15}$$

where $\nu_i = \nu(X_{i1}, X_{i2}, \ldots, X_{iN})$ is the expected value of y from an equation with true βs, $\eta_i = b_o + \sum_{j=1}^{K} b_j X_{ij}$ is the expected value of y from the equation of estimates of βs, and p is K + 1. The term $\sum (\nu_i - \eta_i)^2$ can be represented by SSEB called the sum of squared errors bias. Also, Γ can be defined as the standardized total squared

error

$$\Gamma_p = \frac{SSEB_p}{\sigma^2} + \frac{1}{\sigma^2} \sum_{i=1}^{N} \mathrm{Var}\,(y_{ip}) \tag{16}$$

It is known [Daniel and Wood (Ref 8:86)] that

$$\sum_{i=1}^{N} \mathrm{Var}\,(y_{ip}) = p\,\sigma^2 \tag{17}$$

Combining (16) and (17) forms

$$\Gamma_p = \frac{SSEB_p}{\sigma^2} + p \tag{18}$$

The error sum of squares (SSE) is defined as

$$SSE = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{19}$$

and

$$E(SSE) = \sum_{i=1}^{N} (\nu_i - E(\hat{y}_i))^2 + (N-p)\sigma^2 \tag{20}$$

Because $E(\hat{y}_i) = \eta_i$, then

$$E(SSE) = \sum_{i=1}^{N} (\nu_i - \eta_i)^2 + (N-p)\sigma^2 \tag{21}$$

or

$$E(SSE) = SSEB_p = (N-p)\sigma^2 \tag{22}$$

Combining (18) and (22) gives

$$\Gamma_p = \frac{E(SSE_p)}{\sigma^2} - (N - 2p) \tag{23}$$

Now define C_p as an estimate of Γ_p

$$C_p = \frac{SSE_p}{S^2} - (N - 2p) \tag{24}$$

where s^2 is an estimate of σ^2. Note from (22) that when the correct model is used the bias, SSB_p, goes to zero and C_p goes to p.

Then the objective of the C_p search is to find the p-term equation which has a $\frac{C_p}{p}$ value nearest one and therefore minimum bias. A drawback to the method is that it is sensitive to the S^2 estimate of variance. The C_p obtained from two different models may not be comparable unless the same value of S^2 was used in both. For this reason, the LLSCFP allows the option of using S^2 from an entire set of input variables, from a subset of variables, or a user supplied value. It seems sensible, when a large number of models are being compared, to supply a constant value of S^2 so the C_p values can be compared.

It should be noted that the form of the Γ_p and C_p equations (23) and (24) indicate the importance of eliminating unnecessary variables from a model. The removal of one variable has the capability to remove as much as two units from the standardized squared error.

II. Selection of Variables

Variables

Westinghouse collected data amounting to 63 points on 21 physical and usage characteristics and 6 logistics characteristics of line replaceable avionics units, with the intent to predict logistics characteristics from physical and usage characteristics. Sources used to collect the data were: existing Air Force Data Systems, site visits to Air Force logistics facilities, Westinghouse activities, published reports, and engineering analysis of LRUs. The following paragraphs describe briefly each of the variables. For a more in-depth description of the variables and the manner in which they were collected, see the Westinghouse report (Ref 22:18).

The first six independent variables are measures of physical characteristics. The Unit Price is measured in dollars per LRU. The Volume is measured in cubic feet. Weight is measured in pounds. Component Count is a measure of the number of electrical components of an LRU and does not include mechanical devices, connectors, or structure. Component Density is simply Component Count divided by Volume. While Westinghouse used Component Density in their analysis, it will not be used here, because the matrix of data would be singular in the log-linear model which will eventually be used. Power

Dissipation is measured in input watts minus output watts.

The next five independent variables, measures of component type, are in terms of percentages and are additive to unity. The variable names are descriptive, but if information on the way the measurements are determined is desired, the Westinghouse report (Ref 22:18) should be consulted. The variables are: Fraction Digital, Fraction Analog, Fraction Electromechanical, Fraction Power Supply, and Fraction Transmitter.

The twelfth variable, Fraction Solid State, is a measure of LRU technology, the percentage of components in the LRU that are solid state in nature.

The next data element collected concerned aircraft type and usage. Three types of aircraft were Fighter, Bomber, and Cargo. The three types of usage were Navigation, Sensory, and Communications. In both the Westinghouse and this study these parameters were used as indicators, but in different ways. Westinghouse coded them as follows:

Bomber	1	0
Cargo	0	1
Fighter	0	0
Sensory	1	0
Communications	0	1
Navigation	0	0

In addition, Westinghouse found interactions between certain of the two types useful in their analysis. These were coded as follows:

Bomber Sensory	1	0	0
Bomber Communications	0	1	0
Cargo Communications	0	0	1

In this study a different approach was taken to these variables. Interaction between all of the aircraft types and usages were considered as indicator variables and coded as follows:

Nav Fighter	1	0	0	0	0	0	0
Nav Bomber	0	1	0	0	0	0	0
Nav Cargo	0	0	1	0	0	0	0
Sensor Fighter	0	0	0	1	0	0	0
Sensor Bomber	0	0	0	0	1	0	0
Comm Fighter	0	0	0	0	0	1	0
Comm Bomber	0	0	0	0	0	0	1
Comm Cargo	0	0	0	0	0	0	0

The data base does not include any sensory equipment on Cargo aircraft, thus it is not used as a variable.

The final independent variable identified by Westinghouse is the Percentage of Failures Detected by Built-In-Test (BIT). This variable is intended to be a measure of the effectiveness of the Built-In-Test/ Fault-Isolation-Test (BIT/FIT) capabilities of each LRU.

The six independent variables which Westinghouse identified are Maintenance Manhours/Operating Hour, Mean Time Between Failures, Mean Time Between Maintenance Actions, Logistics Support Cost/ Operating Hour, Training Costs/Operating Hour, and Percentage Not Repairable This Station. The Operating Hours used to normalize is an estimate of the amount of time an LRU is turned on, data which is not recorded. Instead, Westinghouse used flying hours/year of the aircraft using an LRU multiplied by a scaling factor which they estimated

for each of the three types of aircraft. All of the maintenance and operating hours and cost data for these six variables are total yearly figures.

The previous analysis regressed the independent variables against each of the dependent variables individually. Because it is not the purpose of this study to dispute the earlier one, but rather to develop a method for the personnel of the Avionics Laboratory to perform the regression, only one of the dependent variables was chosen for illustration. Logistics Support Cost/Operating Hour was chosen because the Westinghouse report recommended that further study be conducted on it.

Table I contains the variables and the abbreviations used in this and the previous report. An asterisk indicates that the variable was not used in the study indicated. Table II contains the equation found by Westinghouse.

Models

It is well known that the selection of the model is very important to the goodness of fit of a regression and its predictive ability. In light of this, a balance was sought that would yield a model with elements of both simplicity and goodness of fit.

The first model considered was the simplest form of multiple regression equation, namely

Table I

List of Variables

Name	Abbreviations	
	Westinghouse	This Report
Unit Price	UP	UP
Volume	V	V
Weight	W	W
Component Count	CC	CC
Component Density	CD	*
Power Dissipation	PD	PD
Fraction Solid State	FSS	%SS
Fraction Digital	FDI	%DIG
Fraction Analog	FAN	%AN
Fraction Electromechanical	FEM	%EM
Fraction Power Supply	FPS	%PS
Fraction Transmitter	FXR	%XMTR
Bomber	IBOM	*
Cargo	ICAR	*
Sensory	ISEN	*
Communications	ICOM	*
Navigation-Fighter	*	NF
Navigation-Bomber	*	NB
Navigation-Cargo	*	NC

Table I (cont'd)

Name	Abbreviations	
	Westinghouse	This Report
Sensory-Fighter	*	SF
Sensory-Bomber	BOMSEN	SB
Communications-Fighter	*	CF
Communications-Bomber	BOMCOM	CB
Communications-Cargo	CARCOM	COMMC
Fraction BIT/FIT	BIT/FIT	BF
Logistics Support Cost/ Operating Hour	LSC/OH	LSC/OH
Maintenance Manhours/ Operating Hour	MMH/OH	*
Mean Time Between Failures	MTBF	*
Mean Time Between Maintenance Actions	MTBMA	*
Training Cost/Operating Hour	TRAIN/OH	*
Not Repairable This Station	NRTS	*

*Not used in the analysis.

Table II

Logistics Support Cost/Operating Hour Equation
Generated by Westinghouse

$$\ln(LSC/OH) = \alpha_0 + \sum_{i=1}^{21} b_i X_i$$

$R^2 = .8916$ $\quad$ $\bar{R}^2 = .9283$ $\quad$ F-value = 25.3

i	b_i	X_i	Partial F
0	-8.15108		
1	3.86111	(IBOM - .286)	36.0
2	3.66533	(ICAR - .270)	31.4
3	-4.85271×10^{-1}	(ISEN - .254)	3.6
4	-2.56663	(IBOM - .286) (ISEN - .254)	37.2
5	-1.66262	(IBOM - .286) (ICOM - .206)	12.2
6	-7.67253×10^{-1}	(ICAR - .270) (ICOM - .206)	3.2
7	1.27356×10^{-2}	FPS	6.8
8	2.25967×10^{-2}	(FAN - 63.3)	36.0
9	-7.42999×10^{-3}	(FSS - 61.1)	9.0
10	2.38503	(UP - 1.64)	27.0
11	-9.20384×10^{-11}	$(UP - 133606.3)^2$	25.0
12	-1.52864×10^{-4}	$(W - 64.3)^2$	8.4
13	-1.07105×10^{-3}	$(FAN - 48.9)^2$	33.6
14	1.20418×10^{-3}	$(FEM - 47.0)^2$	33.6
15	7.10025×10^{-4}	$(FXR - 40.2)^2$	10.9
16	-1.61651×10^{-4}	$(FSS - 51.9)^2$	2.2

Table II (cont'd)

i	b_i	X_i	Partial F
17	-1.11568×10^{-6}	$(PD - 722)^2$	7.3
18	5.00996	$(UP - 1.68)^2$	42.2
19	1.70042×10^{-3}	$(BF - 27.3)^2$	13.0
20	4.60293×10^{-1}	LN(UP)	31.4
21	2.35583×10^{-1}	LN(V)	4.8

$$y = \alpha_o + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon \qquad (1)$$

When this model was used to form a prediction equation on the original data using SPSS, it did not provide satisfactory goodness of fit, having a Correlation Coefficient Squared of approximately 0.45, indicating that the values predicted by the model were poorly matched to the actual values. The model fell far short of the R^2 of approximately 0.93 achieved by Westinghouse.

A model was desired which would allow more possibility of interactions and not be restricted to linearity of variables. At the same time it was noted that the five measures of component type, namely %DIG, %AN, %EM, %PS, %XMTR, along with the Fraction of Malfunctions Detected by BIT/FIT, BF, could all be converted to indicator form with little loss of information, as almost all of the measurements were near 0% or 100%. This meant that the variable list now

consisted of 13 indicator or dummy variables and 6 ordinary independent variables.

The model settled on was one of considerable complexity but offered many possibilities for interactions between variables. The Product of Powers model is of the form

$$y = e^{(\alpha_o + \sum_i \alpha_i D_i)} \prod_j x_j^{(\beta_{jo} + \sum_i \beta_{ji} D_i)} \tag{25}$$

where α, β, and x have the same meaning as described in the previous model, and D_i is indicator or dummy variable i and j is the index of ordinary variable x. Because this model is not linear, least squares regression can not be used on it without a transformation. If the natural logarithm is taken on both sides of (25), the resulting equation is

$$\ln y = \alpha_o + \sum_{i=1}^{13} \alpha_i D_i + \sum_{j=1}^{6} \beta_{jo} \ln xj$$

$$+ \sum_{j=1}^{6} \sum_{i=1}^{13} \beta_{ji} D_i \ln x_j \tag{26}$$

Linear least squares regression can be applied to the model in (26), but there are now 97 possible variables. Because there are only 63 data points, selection of variables has become of paramount importance. And because only one of the packages used in this study, SPSS, is capable of handling this number of terms, some method was needed to

eliminate terms prior to regression. That is where a test of equality of regression populations was useful.

Test of Equality of Regression Populations

The coefficients and constant terms, βs and αs in (26) could be written in vector form as

$$\underline{\beta} = \begin{bmatrix} \alpha_0 + \alpha_1 + \ldots + \alpha_{13} \\ \beta_{1\ 0} + \beta_{1\ 1} + \ldots + \beta_{1\ 13} \\ \beta_{2\ 0} + \beta_{2\ 1} + \ldots + \beta_{2\ 13} \\ \cdot \\ \cdot \\ \cdot \\ \beta_{6\ 0} + \beta_{6\ 1} + \ldots + \beta_{6\ 13} \end{bmatrix} \tag{27}$$

by noting that $\ln x_j$ is a common term in the second and third term of (26) and that $D_i = 0$ or 1. If only one indicator is considered at a time and all others are considered constant, (27) becomes

$$\underline{\beta}_b = \begin{bmatrix} a_o + a_i \\ \beta_{1\ 0} + \beta_{1\ i} \\ \beta_{2\ 0} + \beta_{2\ i} \\ \cdot \\ \cdot \\ \cdot \\ \beta_{6\ 0} + \beta_{6\ i} \end{bmatrix} \tag{28}$$

Equation (28) can be divided into two subsets depending on whether $D_i = 0$ or 1.

$$
\underline{\beta}_a = \begin{bmatrix} \alpha_0 \\ \beta_{10} \\ \beta_{20} \\ \cdot \\ \cdot \\ \cdot \\ \beta_{60} \end{bmatrix} \tag{29}
$$

if $D_i = 0$. If $D_i = 1$, then equation (28) is the case.

If it could be shown that equation (28) and (29) are not significantly different, then (29) would be used and the interaction terms between that D_i and all x_j as well as that $\alpha_i D_i$ would not be included in the model. Further, if it could be shown that $\underline{\beta}_a = \underline{\beta}_b$ where

$$
\underline{\beta}_a = \begin{bmatrix} \alpha_0 + \alpha_i \\ \beta_{10} + \beta_{1i} \\ \beta_{20} \\ \beta_{30} + \beta_{3i} \\ \cdot \\ \cdot \\ \cdot \\ \beta_{60} + \beta_{6i} \end{bmatrix} \tag{30}
$$

and $\underline{\beta}_b$ = equation (28), then the individual interaction between x_2 and that D_i would not be needed in the model. Any of the α_i or β_{ji} terms could be so tested either singly as above or in combinations.

Chow describes such a test (Ref 7:599) and calls it a Test of Equality Between Subsets of Coefficients in Two Regressions. The same technique was described by Fisher (Ref 12:364).

Under the alternative hypothesis $\underline{\beta}_a \neq \underline{\beta}_b$ the model becomes

$$y_1 = X_1 \beta_1 + \epsilon_1 = Z_1 \gamma_1 + W_1 \delta_1 + \epsilon_1 \tag{31}$$

$$y_2 = X_2 \beta_2 + \epsilon_2 = Z_2 \gamma_2 + W_2 \delta_2 + \epsilon_2 \tag{32}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} Z_1 & 0 & W_1 & 0 \\ 0 & Z_2 & 0 & W_2 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{33}$$

where Z_1, Z_2, W_1, and W_2 are submatrices of the X matrix of data

$$X = \begin{bmatrix} 1 & X_{1\,1} & X_{1\,2} & \cdots & X_{1\,K} \\ 1 & X_{2\,1} & X_{2\,2} & \cdots & X_{2\,K} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ 1 & X_{N\,1} & \cdots & & X_{N\,K} \end{bmatrix} \tag{4}$$

The matrix Z_1 contains those elements of X for the variables which are being tested and in which the D_i being considered is equal to zero. Z_2 contains those elements of X for the variables which are being tested and in which the D_i being considered is equal to one. W_1 contains those elements of X for the variables which are not being tested and in which the D_i equals zero. W_2 contains the remainder of X, those elements for which the variable is not being tested and D_i equals one. The vector γ_1 contains the regression coefficients which are being tested, assuming D_i equals zero, such as $\beta_{2\,0}$ in (30). The

vector γ_2 contains the coefficients being tested assuming D_i equals one such as $\beta_{2\,0} + \beta_{2\,i}$. The vector δ_1 contains the regression coefficients not being tested assuming D_i equals zero. The vector δ_2 contains the coefficients not being tested assuming D_i equals one. For example, if it is desired to test whether

$$\begin{bmatrix} \alpha_o + \alpha_1 \\ \beta_{1\,0} + \beta_{1\,1} \\ \beta_{2\,0} \\ \beta_{3\,0} + \beta_{3\,1} \\ \beta_{4\,0} + \beta_{4\,1} \\ \beta_{5\,0} + \beta_{5\,1} \\ \beta_{6\,0} + \beta_{6\,1} \end{bmatrix} = \begin{bmatrix} \alpha_o + \alpha_1 \\ \beta_{1\,0} + \beta_{1\,1} \\ \beta_{2\,0} + \beta_{2\,1} \\ \beta_{3\,0} + \beta_{3\,1} \\ \beta_{4\,0} + \beta_{4\,1} \\ \beta_{5\,0} + \beta_{5\,1} \\ \beta_{6\,0} + \beta_{6\,1} \end{bmatrix} \tag{34}$$

then

$$Z_1 = \begin{bmatrix} X_{1\,2} \\ X_{2\,2} \\ \cdot \\ \cdot \\ \cdot \\ X_{n\,2} \end{bmatrix} \tag{35}$$

and

$$Z_2 = \begin{bmatrix} X_{n+1\,2} \\ X_{n+2\,2} \\ \cdot \\ \cdot \\ \cdot \\ X_{n+m\,2} \end{bmatrix} \tag{36}$$

Assuming that D_1 equals zero for the first n observations and one for the next m observation.

$$W_1 = \begin{bmatrix} 1 & x_{11} & x_{13} & x_{14} & x_{15} & x_{16} \\ 1 & x_{21} & x_{23} & x_{24} & x_{25} & x_{26} \\ \vdots & \vdots & \vdots & & & \vdots \\ 1 & x_{n1} & x_{n3} & x_{n4} & x_{n5} & x_{n6} \end{bmatrix} \tag{37}$$

and

$$W_2 = \begin{bmatrix} 1 & x_{n+1\,1} & x_{n+1\,3} & x_{n+1\,4} & x_{n+1\,5} & x_{n+1\,6} \\ 1 & x_{n+2\,1} & x_{n+2\,3} & x_{n+2\,4} & x_{n+2\,5} & x_{n+2\,6} \\ \vdots & \vdots & & & & \vdots \\ 1 & x_{n+m\,1} & x_{n+m\,3} & x_{n+m\,4} & x_{n+m\,5} & x_{n+m\,6} \end{bmatrix} \tag{38}$$

The vectors would be

$$\gamma_1 = \begin{bmatrix} \beta_{20} \end{bmatrix} \tag{39}$$

$$\gamma_2 = \begin{bmatrix} \beta_{20} + \beta_{21} \end{bmatrix} \tag{40}$$

$$\delta_1 = \begin{bmatrix} \alpha_0 \\ \beta_{10} \\ \beta_{30} \\ \beta_{40} \\ \beta_{50} \\ \beta_{60} \end{bmatrix} \tag{41}$$

$$\delta_2 = \begin{bmatrix} a_o + a_1 \\ \beta_{1\,0} + \beta_{1\,1} \\ \beta_{3\,0} + \beta_{3\,1} \\ \beta_{4\,0} + \beta_{4\,1} \\ \beta_{5\,0} + \beta_{5\,1} \\ \beta_{6\,0} + \beta_{6\,1} \end{bmatrix} \tag{42}$$

Under the null hypothesis $\underline{\beta}_a = \underline{\beta}_b$, equation (34), the model becomes

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} Z_1 & W_1 & 0 \\ Z_2 & 0 & W_2 \end{bmatrix} \begin{bmatrix} \gamma \\ \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{43}$$

where Z_1, Z_2, W_1, W_2, δ_1, and δ_2 are the same as in (33) and contains the regression coefficients of the variables being tested assuming D_i equals zero. In the example being used here,

$$\gamma = \beta_{2\,0} \tag{44}$$

Under the null hypothesis, the least squares estimators of γ, δ_1, and δ_2 are

$$\begin{bmatrix} c_o \\ d_{1\,0} \\ d_{2\,0} \end{bmatrix} = \begin{bmatrix} Z_1'Z_1 + Z_2'Z_2 & Z_1'W_1 & Z_2'W_2 \\ W_1'Z_1 & W_1'W_1 & 0 \\ W_2'Z_2 & 0 & W_2'W_2 \end{bmatrix}^{-1} \begin{bmatrix} Z_1' & Z_2' \\ W_1' & 0 \\ 0 & W_2' \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{45}$$

where y_1 is a vector of observed y values for which the corresponding D_i equals zero, and y_2 is a vector of observed y values for which the corresponding D_i equals one. Continuing the same example as

previously,

$$\underline{y}_1 = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} \tag{46}$$

and

$$\underline{y}_2 = \begin{bmatrix} y_{n+1} \\ y_{n+2} \\ \cdot \\ \cdot \\ \cdot \\ y_{n+m} \end{bmatrix} \tag{47}$$

Under the alternative hypothesis,

$$\begin{bmatrix} c_1 \\ c_2 \\ d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} Z_1 Z_1 \\ 0 \\ W_1 Z_1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & Z_1 W_1 & 0 \\ Z_2 Z_2 & 0 & Z_2 W_2 \\ 0 & W_1 W_1 & 0 \\ W_2 Z_2 & 0 & W_2 W_2 \end{bmatrix} \cdot \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \\ W_1 & 0 \\ 0 & W_2 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{48}$$

Once the estimated coefficients are found, an F-test can be used to test the null hypothesis. If $m > p$, where p is the length of the vectors in (34) the test is

$$F_{(q,\, m+n-2p)} = \frac{\| Z_1 c_1 + W_1 d_1 - Z_1 c_o - W_1 d_{10} \|^2 + \| Z_2 c_2 + W_2 d_2 - Z_2 c_o - W_2 d_{20} \|^2}{\| y_1 - Z_1 c_1 - W_1 d_1 \|^2 + \| y_2 - Z_2 c_2 - W_2 d_2 \|^2}$$

$$\cdot \frac{m+n-2p}{q} \tag{49}$$

If $p-q \leq m \leq p$ where q is the number of variables being tested or the length of vector γ_1 in (39), the test is

$$F_{(m-p+q,\, n-p)} = \frac{\|Z_1c_1+W_1d_1-Z_1c_o-W_1d_{10}\|^2 + \|y_2-Z_2c_o-W_2d_{20}\|^2}{\|y_1-Z_1c_1-W_1d_1\|^2} \cdot \frac{n-p}{m-p+q} \qquad (50)$$

If the calculated F value is greater than the table F value at the desired level of confidence, then reject H_0 and include the interaction term in the model. If the calculated F is lower than the table F, accept H_0 and do not include the interaction term in the model. If $m < p-q$, the test can not be performed in which case the null hypothesis has not been rejected and the variables are not included in the model.

In doing this study, the Chow test was performed on each $\beta_{j0} + \beta_{ji}$ and each $\alpha_o + \alpha_i$ term individually for each combination of j and i. Thus the test had the potential to eliminate any of 78 β_{ji} and 13 α_i terms of the 97 possible in the product of powers model used.

To perform the calculations for these tests three programs were written. The first program called YSEPR simply separated the observed values of y into the vectors $\underline{y}_1$ and $\underline{y}_2$ for each of the 13 dummy variables. A listing of the program is contained in the appendix on Figure 4.

The second program, called SUMS, was written to calculate

all of the combinations of sums of variables, sums of squared variables, and sums of cross products of variables grouped as in YSEPR by the value of each dummy variable. These sums are calculated because when the multiplications required in the large matrices in (45) and (48) are performed, the result would be some of those elements calculated by SUMS. Rather than perform the calculations repeatedly when only a small systematic change is needed for each test, they are calculated only once by SUMS and read when needed by the third program. A listing of SUMS is contained in the appendix on Figure 5.

The third program, called CHOW, performs the calculations in (34), (37), (38), and (39) and outputs the F values with the parameters necessary to test the hypotheses. A listing of CHOW is contained in the appendix on Figure 6. Table III contains the calculated F values generated by CHOW. Table IV contains the ranks of the F values generated.

In the tests performed here, $p-q \leq m \leq p$ never occurred. Therefore, the degrees of freedom for all of the tests were q and $m+n-2p$ or 1 and 49 where $m+n = 63$ and $p = 7$.

The coefficients associated with dummy variables SB, CF, CB, and PS could not be tested because m was less than p-q. As a result, there were 91-24 = 67 variables tested using the Chow test. If the overall hypothesis that none of the $\beta_{j0} + \beta_{ji}$ terms differ from β_{j0} or $\alpha_0 + \alpha_i$ terms differ from α_0 is true, then the probability that at

least one of the individual hypotheses will be rejected would be less than α^* where

$$\alpha^* = 1 - (1 - \alpha)^{67} \tag{51}$$

and is the level at which individual tests are conducted. Then

$$\alpha = 1 - (1 - \alpha^*)^{1/67} \tag{52}$$

If $\alpha^* = 0.10$ is desired then the associated value of α would be 0.00157. Using the tabled F-value for $\alpha = 0.001$ assures that α^* is less than 0.10. Values of α^* less than 0.10 can not be tested at this time because available tables of F-values only go down to 0.001. When the calculated F is larger than $F_{0.001, 1, 49} = 12.11$ the individual hypothesis is rejected and the interaction β_{ji} term can not be removed from the model on the basis of this test. Table III contains the calculated F-values for each of the tested terms. Variable Ø indicates the α_{0i} term. The numbering system used in program CHOW is shown at the bottom of the table.

Only 13 terms had F-values lower than the critical F. These in addition to the 24 variables which could not be tested and therefore did not fail a test of the null hypothesis mean that 37 variables have been eliminated from the 97 started with, leaving 60 variables still in the model. While this appears disappointing at first glance, leaving more variables still in the model than LLSCFP or Leaps and Bounds can

Table III

Calculated F Values from CHOW Test

Dummy	Variable 0*	1	2	3	4	5	6
1	46.792	35.957	25.792	25.083	25.321	4.445	30.243
2	48.9942	31.165	27.368	17.834	37.558	2.207	13.168
3	47.188	35.020	20.261	42.983	.434	1.664	2.142
4	25.038	4.028	43.242	42.912	35.336	12.666	.138
5	NT	NT	NT	NT	NT	NT	NT
6	NT	NT	NT	NT	NT	NT	NT
7	NT	NT	NT	NT	NT	NT	NT
8	49.002	36.238	36.429	47.934	44.038	48.998	35.319
9	49.440	40.300	14.671	45.948	.735	7.368	.681
10	48.900	35.818	43.052	21.842	28.715	7.510	19.331
11	NT	NT	NT	NT	NT	NT	NT
12	54.984	42.913	45.636	35.172	15.782	17.991	15.147
13	46.648	39.791	22.918	39.312	11.510	15.345	11.583

Variable Names

1	2	3	4	5	6
UP	V	W	CC	%SS	PD

Dummy Names

1	2	3	4	5	6	7	8	9	10	11	12	13
NF	NB	NC	SF	SB	CF	CB	DIG	AN	EM	PS	XMTR	BF

NT indicates insufficient points in a subset to test.

*Variable Ø indicates the a_iD_i term.

handle, experimentation will be conducted in later chapters in which the rankings of the F-values will be used to try to preselect a set of variables for a model. The rankings of the F-values are shown in Table IV, with the lowest F-value being ranked 1. Also sub-optimized solutions, in which the 60 variables are broken into subsets to search for smaller subsets that can be combined, to form manageable sets, will be tried. Table V contains a list of the variables remaining in the model after the Chow test.

Table IV

Rank of F-Values from CHOW Test

Dummy	Variable 0*	1	2	3	4	5	6
1	55	39	29	27	28	9	32
2	59	33	30	20	42	7	15
3	56	34	23	48	2	5	6
4	26	8	50	46	37	14	1
5	NT	NT	NT	NT	NT	NT	NT
6	NT	NT	NT	NT	NT	NT	NT
7	NT	NT	NT	NT	NT	NT	NT
8	61	40	41	57	51	60	36
9	62	45	16	53	4	10	3
10	58	38	49	24	31	11	22
11	NT	NT	NT	NT	NT	NT	NT
12	63	47	52	35	19	21	17
13	54	44	25	43	12	18	13

NT indicates that there were insufficient points in one of the subsets to perform the test.

*Variable Ø indicates the $\alpha_i D_i$ term.

Table V

Variable Remaining in the Model After the Chow Test

1 UP	21 NF*V	41 DIG*%SS
2 V	22 NF*W	42 DIG*PD
3 W	23 NF*CC	43 AN*UP
4 CC	24 NF*PD	44 AN*V
5 %SS	25 NB*UP	45 AN*W
6 PD	26 NB*V	46 EM*UP
7 NF	27 NB*W	47 EM*V
8 NB	28 NB*CC	48 EM*W
9 NC	29 NB*PD	49 EM*CC
10 SF	30 NC*UP	50 EM*PD
11 SB	31 NC*V	51 XMTR*UP
12 CF	32 NC*W	52 XMTR*V
13 CB	33 SF*V	53 XMTR*W
14 DIG	34 SF*W	54 XMTR*CC
15 AN	35 SF*CC	55 XMTR*%SS
16 EM	36 SF*%SS	56 XMTR*PD
17 PS	37 DIG*UP	57 BF*UP
18 XMTR	38 DIG*V	58 BF*V
19 BF	39 DIG*W	59 BF*W
20 NF*UP	40 DIG*CC	60 BF*%SS
		61 LSC/OH (Dependent Variable)

III. Stepwise Regression

Mechanics

The three most commonly used iterative techniques for determining the proper variables in a regression are; backward elimination, forward selection, and stepwise regression.

All three of the above methods make use of the Partial F-test. In this test, the explained sums of squares (SSR) are decomposed into components attributable to each independent variable.

In the standard regression method of decomposition, each variable is treated as if it had been added to the regression equation in a separate step after all other variables had been included. These F-values are then used to determine the next variable to enter or leave the equation, depending on the type of iterative regression being performed. The F-value is given by

$$F = \frac{\text{SSR due to } x_i/1}{\text{SSE}/(N-K-1)}$$

(Ref 20:336)

$$= \frac{r^2_{y(i.1,2,\ldots K)}/1}{1-R^2_{y.1,2,\ldots K}/(N-K-1)} \tag{53}$$

The term $r^2_{y(1.1,2,\ldots K)}$ is the part correlation indicating the relationship between the observed y and the residual of independent variable i

from which the effects of the other independent variables have been removed.

If the hierarchical decomposition method is used instead of the standard regression method, the order of inclusion must be specified and is used to determine the order to enter variables rather than a partial F test. The variable included first, the one with the highest assigned inclusion level, is evaluated by the ratio,

$$F = \frac{r^2_{y1}/1}{(1 - R^2_{y.1,2,\ldots,K})\ (N-K-1)} \qquad (54$$

The second regression coefficient is tested by the ratio,

$$F = \frac{r^2_{y(2.1)}/1}{(1 - R_{y.1,2,\ldots,K})/(N-K-1)} \qquad (55)$$

Ref 20:337)

$$= \frac{\text{incremental SS due to } x_2/1}{SSE/(N-K-1)}$$

Each successive variable to be included in a hierarchical fashion would be evaluated in the same manner as indicated above, with a squared part correlation of the form $r^2_{y(i.1,2,\ldots,i-1)}$, where i is the variable being added, in the numerator. All of the ratios above should be compared to the tabled F distribution with 1 and (N-K-1) degrees of freedom.

The hierarchical method of regression is used when there is some basis for believing that some variables will explain more variance than others before the regression is accomplished. Because there were no prior feelings about the variables used in this analysis, the hierarchical method was not used.

In the background elimination method, a regression equation with all of the possible terms is the starting point. Then partial F-value is calculated for each variable. If the lowest partial F-value is less than some preselected F-value, the variable corresponding to the value is removed from consideration, and the procedure is repeated from calculation of F-values for the new equation. If at any iteration there are no variables that have an F-value low enough to cause removal, the equation is adopted as calculated.

The forward selection method operates by entering variables one at a time until a satisfactory equation is reached. The order of inclusion is determined by using partial correlation coefficients as a measure of relative importance of the variables not yet in the equation. At each step, the variable with the highest partial correlation coefficient, that is, the variable with the highest correlation to the dependent variable after allowing for the effect of previously included variables, is brought into the equation. This procedure is repeated until the partial F-value of the latest entered term is less than some preselected value; then the equation is adopted.

The stepwise regression method is a refinement of the forward selection method in that at every stage of the procedure, the variables included in the model in earlier stages are examined. Thus a variable which was entered at an earlier stage but has been rendered unimportant by its relationship to later entered variables will be detected and removed from the equation. Again, this is done by comparing the partial F-values to a preselected F-value to find any variables to be removed. In later iterations, the removed variable is treated the same as a variable that has never entered the equation. The criterion for inclusion into the stepwise model is the same as in the forward selection method.

While calculation of the partial F-values needed to use these regression methods would be too time consuming to perform manually, both of the regression packages used in this chapter, SPSS and BMD, provide them, making the two packages easy and convenient to use. Because stepwise regression incorporates the main features of backward elimination and forward selection, it is considered to be the best of the three methods (Ref 11:172). While it has been shown that the three methods do not always choose the same subset (Ref 17:9), the stepwise method is more likely to choose the best. For that reason, of the three methods, stepwise regression will be used here.

In using iterative regression method, two pitfalls should be kept in mind. First, no significance should be attached to the order in which

variables are entered or removed from the model. The first variable entered is not necessarily the most important one when other variables have been entered. The partial F-value of each variable must be considered for each variable after each step to determine relative significance. Secondly, there is no guarantee that any of the three methods will supply an optimal equation, because of the restriction of removing and entering one variable at a time.

It was previously stated that preselected values of the F distribution are used as criterion in including and deleting variables from the equation. In both the SPSS and BMD packages these values are called FIN and FOUT respectively. A problem arises in preselecting these values, as the degrees of freedom are not known exactly before a regression is run. The tabled F-value has degrees of freedom 1 and (N-K-1) where N is the number of data points and K is the number of independent variables in the equation. As the primary goal of an analyst is normally to maximize some measure of goodness-of-fit, such as R, while a secondary goal is to minimize K, K is not known exactly before the regression. Therefore, the values of FIN and FOUT must be determined from an estimated K. As a result some experimentation may be necessary to obtain the value of K which yields desired values for the goodness-of-fit measure.

An alternative is to use the default F-values for FIN and FOUT, 0.01 and 0.005, respectively. These values are designed to include

all variables and not remove any from the equation; in effect, a forced fit unless there are a large number of variables. In the course of the analysis, regressions were performed on identical data using varying values for FIN and FOUT. When values other than default were used, truncated versions of the same equation obtained from the default values were identified. It appears the simplest procedure may be to use default values of FIN and FOUT until the value of K is known, then use the table value of F at degrees of freedom 1 and (N-K-1) as FIN. FOUT should be slightly smaller than FIN, but other than that its selection is rather arbitrary.

The Packages

As would be expected when using packages that operate in the same way, SPSS and BMD yield the same answers and even employ the same format of output. Each prints at each step, the value of the multiple correlation coefficient, R, the standard error of estimate, an analysis of variance table, the coefficients of all variables currently in the equation, the standard error of the coefficients, and the partial F-value of all independent variables whether or not in the equation. In addition, SPSS provides R^2 and $\bar{R}^2$. Both also list the value of calculated F to test the hypothesis

$$
H_0: \begin{bmatrix} a \\ \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_K \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \tag{56}
$$

Plots of residuals versus the sequence of cases in a file can be selected as options in both packages. Additionally, BMD will make plots of residuals versus all or selected independent variables as an option. Other options for both input and output are available and identified in the appropriate manuals (Ref 20:352) and (Ref 10:235). Examples of output are available in both manuals (Ref 20:360) and (Ref 10:249) and partial output prepared for this report are included in the appendix in Figures 7 and 8.

SPSS is capable of handling up to 100 independent variables per fit and BMD can handle up to 80. While both of the packages have the capability to compute transformations, the feature was not used. Instead, all interactions were computed using a short FORTRAN program and stored with the original variables on a permanent storage file to avoid recalculating them numerous times. Table VI indicates the location of all variables on the tape used for both stepwise regression and Leaps and Bounds.

Table VI

Numbers of the Variables on Permanent File Used in SPSS and Leaps and Bounds

These are the variable numbers found on the SPSS output.

1	2	3	4	5	6	7	8	9	10	11	12	13	14
UP	V	W	CC	%SS	PD	NF	NB	NC	SF	SB	CF	CB	DIG

15	16	17	18	19
AN	EM	PS	XMTR	BF

	UP	V	W	CC	%SS	PD
NF	20	21	22	23	24	25
NB	26	27	28	29	30	31
NC	32	33	34	35	36	37
SF	38	39	40	41	42	43
SB	44	45	46	47	48	49
CF	50	51	52	53	54	55
CB	56	57	58	59	60	61
DIG	62	63	64	65	66	67
AN	68	69	70	71	72	73
EM	74	75	76	77	78	79
PS	80	81	82	83	84	85
XMTR	86	87	88	89	90	91
BF	92	93	94	95	96	97
			98 LSC/OH			

Results

First stepwise regression was used on the 60 variables remaining after the Chow test. Table VII provides a summary of the independent variable included by each step of the regression along with R^2 and $\bar{R}^2$ for the first 48 steps. It can be seen that step 25 shows a jump of almost 0.03 in R^2 and more than 0.04 in $\bar{R}^2$ over the previous step. Steps that follow provide only small increases in R^2 and $\bar{R}^2$. The adjusted R^2 hits a peak at step 45 but actually changes very little between steps 25 and 45. After step 45, $\bar{R}^2$ shows a general downward trend except when a variable is removed. As a result the equation at step 25 was chosen as the best compromise between correlation and number of variables included. The equation has 23 variables, which is two more than the Westinghouse analysis, but it also provides an R^2 of more than 0.02 higher and $\bar{R}^2$ of about 0.03 higher than they achieved. Table VIII contains the coefficients and the partial F-values for each variable in the equation.

Next, as a comparison, a stepwise regression was performed with all 97 variables as inputs, regardless of results of the Chow test. The first six variables selected were the same as in the previous run, but from that point the two equations diverged. The regression with all 97 variables consistently had an R^2 of 0.02 to 0.03 lower than the previous run for the same number of variables included. This could be taken to mean that the variables deleted by the Chow test were

Table VII

Sequence of Stepwise Regression on 60 Variables
Remaining After Chow

Step	Variable	R^2	$\bar{R}^2$
1	W	.57312	.56612
2	AN*UP	.65870	.64733
3	SB	.72117	.70700
4	DIG*W	.76054	.74402
5	NB	.78460	.76571
6	XMTR*CC	.80354	.78249
7	DIG*V	.81802	.79486
8	BF*%SS	.82905	.80372
9	NB*V	.83850	.81107
10	AN	.84931	.82034
11	EM	.86207	.83232
12	PS	.87267	.84211
13	BF*W	.87661	.84387
14	SF	.88006	.84508
15	UP	.88449	.84822
16	%SS	.88799	.84903
17	XMTR*%SS	.89475	.85499
18	NB*W	.90005	.85917
19	AN(removed)	.90005	.86229
20	EM*V	.90507	.86624

Table VII (cont'd)

Step	Variable	R^2	$\bar{R}^2$
21	NB*UP	.90861	.86822
22	AN*W	.91318	.87184
23	DIG	.91814	.87621
24	NF*CC	.92491	.88361
25	NF*UP	.95212	.92388
26	NC	.95429	.92543
27	CC	.95678	.92757
28	DIG*CC	.95891	.92923
29	NF*PD	.96101	.93094
30	SF*%SS	.96283	.93223
31	SF*W	.96431	.93294
32	SF*CC	.96921	.94035
33	NF	.97111	.94221
34	BF*UP	.97226	.94268
35	XMTR*PD	.97351	.94220
36	PD	.97482	.94228
37	EM*CC	.97564	.94261
38	V	.97703	.94217
39	BF	.97821	.94106
40	EM*PD	.97886	.94056
41	CC(removed)	.97886	.94294

Table VII (cont'd)

Step	Variable	R^2	$\bar{R}^2$
42	CF	.97918	.94218
43	EM*UP	.97975	.94129
44	BF*W(removed)	.98027	.94373
45	NC*V	.98080	.94530
46	EM*W	.98189	.94415
47	XMTR*UP	.98221	.94227
48	XMTR	.98101	.94112

indeed not needed and are only clouding the issue when they are allowed in the set being considered.

Finally, an SPSS run was made on the 35 variables that had the highest F-values from the Chow test in an attempt to determine whether the rank of the F-values had any significance. Even with all 35 variables in the equation the R^2 was only 0.88 and $\bar{R}^2$ was only 0.73. This indicates that the variables with the highest F-values are not necessarily the most important in a regression. The ranks of F-values should not be given any significance.

Conclusion

Both SPSS and BMD are very easy to use and require little foreknowledge of the basis of regression. Therein also lies a danger that

Table VIII

Coefficients of the SPSS Regression Equation

$R^2 = 0.95212$ $\bar{R}^2 = 0.92388$ $F = 33.72$

$$\ln(LSC/OH) = \alpha_0 + \sum_i \alpha_i D_i + \sum \beta_{j0} \ln x_j + \sum_j \sum_i \beta_{ji} D_i \ln x_j$$

Variable No.	Variable Name	Coefficient	Partial F
1	UP	0.402702	13.63
3	W	0.084548	0.10
5	%SS	0.412407	37.28
8	NB	11.320694	23.80
10	SF	-1.135445	17.68
11	SB	-1.457859	26.48
14	DIG	3.710527	7.25
16	EM	-2.950970	9.44
17	PS	-0.092716	0.09
20	NF*UP	0.322015	0.07
23	NF*CC	-0.568085	27.14
26	NB*UP	-0.729848	7.51
27	NB*V	-1.803242	9.46
28	NB*W	2.506829	12.27
63	DIG*V	-1.995969	18.20
64	DIG*W	3.034970	17.51
68	AN*UP	-0.272142	7.44

Table VIII (cont'd)

Variable No.	Variable Name	Coefficient	Partial F
70	AN*W	0.758240	8.11
75	EM*V	0.422377	8.71
89	XMTR*CC	0.294839	25.70
90	XMTR*%SS	-0.456146	24.86
94	BF*W	0.697895	25.90
96	BF*%SS	-0.642736	43.88
Constant		-5.315378	79.01

the packages will be misapplied and their output applied blindly. Both manuals (Ref 20:320) and (Ref 10:215) contain an introduction to regression which should be sufficient background for most cases. Of the two manuals, the SPSS is more detailed and easier to follow.

IV. All Possible Regressions

The number of possible subsets given K possible variables increases at the rate of 2^K as K increases, and the number of calculations required to invert the moments matrix for each subset is of the order K^3. By taking advantage of the symmetry of moments matrices and the deletion of unneeded rows and columns as successive regressions are calculated, as well as storing moments matrices for later modification and use, the order of calculations for each subset is brought down to K^2. If only the regression coefficients, their variances, and the sum of squared errors, are wanted, the calculations required are of order K for each subset. If only the sum of squared errors is needed, the number of calculations is less than six per subset (Ref 13:500). But even this would mean 9.5×10^{29} calculations for 97 variables and 6.4×10^9 calculations for only 30 variables just to calculate only the sums of squared errors for each possible regression. Clearly some way to eliminate some possibilities before calculating is needed to make the method feasible for problems of this magnitude.

This is what has been developed by Furnival and Wilson (Ref 13). Their technique, called Leaps and Bounds, is based on the fundamental fact that $SSE(A) \leq SSE(B)$ where A is any set of independent variables

and B is any subset of A. It is impossible for any subset of A to have a lower error sum of squares than A. From this, we can use the SSE of A as a lower bound of the subsets (in the nomenclature of Furnival and Wilson, offspring of A). It is known that regressions of less variables than A that are not offspring of A have a lower SSE than A, then there is no need to investigate the offspring of A as possible optimal solutions. In this way, the number of regressions and calculations are reduced. The amount by which they are reduced is determined by how early in the branching process good lower bounds are found. If a good lower bound is located early, most of the regressions are eliminated from consideration before calculations are performed on them. It was noted in performing this study that the amount of execution time, and hence calculations, varied over a wide range for a given number of variables input. For instance, for the case of 29 independent variables, execution time ranged from 14 to 110 seconds, depending on the variables input.

To illustrate the method, an example from Furnival and Wilson (Ref 13:506) will be duplicated below in Figure 1. There are five independent variables numbered 1 to 5. Underlined variables are those that are going to be removed in offspring equations. Missing variables indicate those that have already been removed from the branch. The numbers in parentheses are the SSE values for the corresponding equation. For example, . $12\underline{45}$ indicates the subset of variables containing

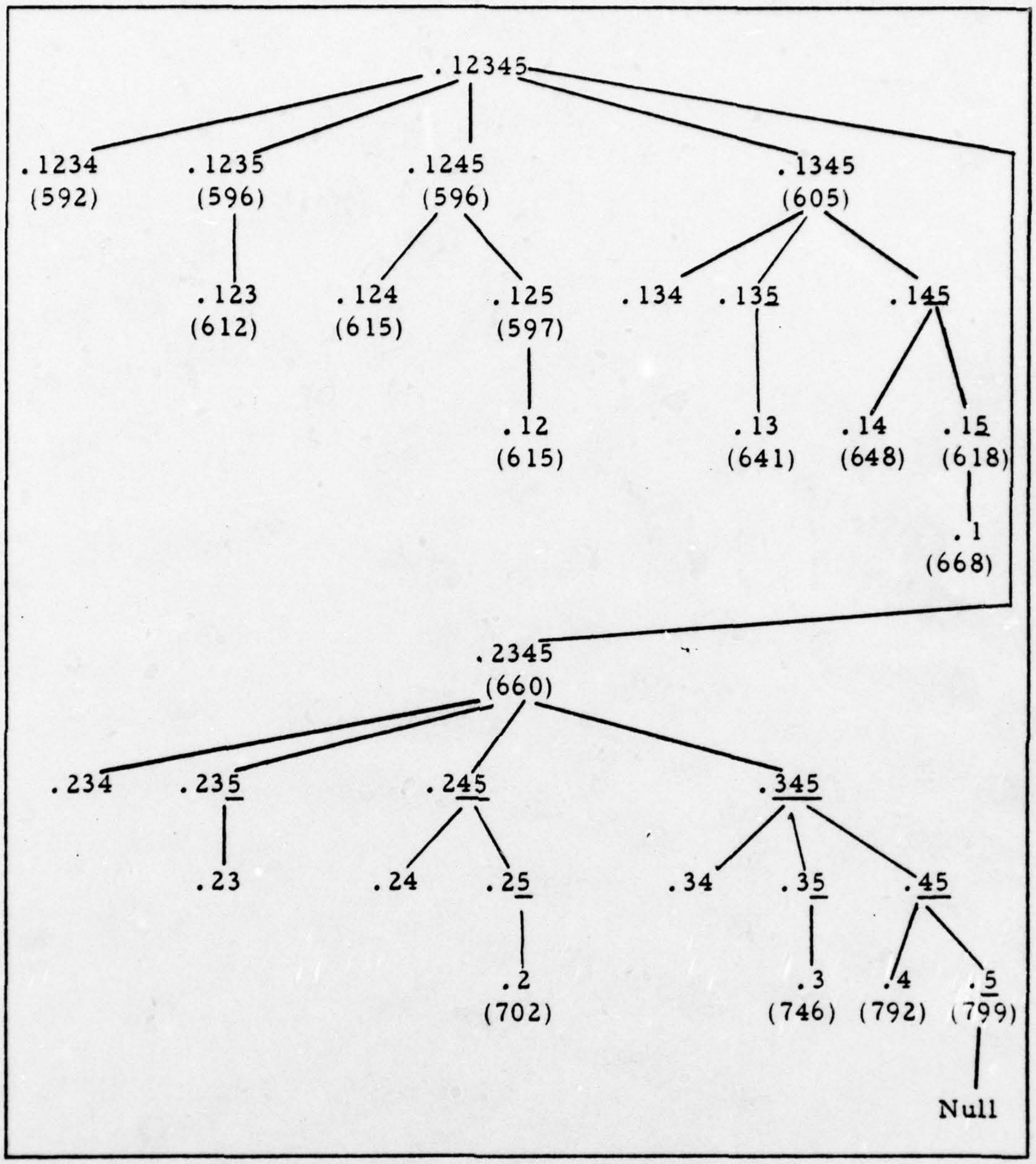


Figure 1. Tree Diagram

1, 2, 4, and 5, from which 4 and 5 will be removed in offspring branches.

There are various methods by which the tree can be traversed;

horizontally, vertically or a combination of those. The simplest for illustrative purposes is the horizontal method whereby equations of the same number of variables are considered together. The process begins by calculating the SSE for each of the four-variable equations. Then the SSE of equation .123 is calculated as 612. If 612 were lower than any of the SSEs from the four variable equations, no three-variable offspring of that four-variable equation would be considered or calculated. Note that equation .12$\underline{5}$ has an SSE of 597, which is lower than the SSE of 605 for equation .13$\underline{45}$ or the SSE of 660 for equation .$\underline{2345}$. Therefore, there is no need to calculate the SSE for equations .134, .13$\underline{5}$, .1$\underline{45}$, .234, .23$\underline{5}$, .2$\underline{45}$, or .$\underline{345}$.

Next begin to evaluate the two-variable equations and compare them to the three and four-variable equations not already eliminated. Note that the SSE of equation .12, 615, is less than that for equation .$\underline{2345}$. Therefore, there is no need to consider two-variable offspring of equation .$\underline{2345}$. We need calculate the SSE only for equations .12, .13, .14, and .1$\underline{5}$.

Finally, begin to evaluate the one-variable models and compare them to the two, three, and four-variable models not already eliminated. If an SSE of a one-variable model is found to be less than that of a two, three, or four-variable equation, then their one-variable offspring need not be considered. In this case, all of the one-variable equations must be considered, and equation .1 is found to be the best

one-variable model.

In this example, only 17 values of SSE were calculated to guarantee that the best one, two, three, and four-variable equations were found. While the traverse used here is not as efficient as that used in the version of Leaps and Bounds in the International Mathematical and Statistical Libraries (IMSL), it is sufficient for illustrative purposes. The IMSL subroutine, called RLEAP, provides as output the m best of each of the 1 to K-variables subsets, where the user supplies the value of m. The output can be based on any or each of the criterion of R^2, adjusted R^2 of C_p. The LLSCFP also uses a search technique based on Cp, but the version available for this study allowed only 29 variables to be input at a time and of these, only 12 at a time could be searched. This would cause the problem to become too fragmented for the number of variables that had to be considered. As a result, the LLSCFP was not used.

Application

It was found that the amount of execution time used by RLEAP grew very quickly above 20 variables. At 20 variables execution time was only a few seconds on the CDC 6600 series computer. But at 29 variables, the execution time could go as high as 138 seconds octal, at 35 variables over 1,000 seconds octal would be required. As a result 29 variables were settled on as a practice maximum input to RLEAP for this study.

Because there were 60 variables remaining from the Chow test, it was necessary to consider them in groups of no more than 29 and then to combine the optimum answers from each group into a new set. The new subset formed in this way was then run through RLEAP to pick the optimum subset from it. The result should then be a relatively small number of variables which can be used as a core to rotate the remaining variables through RLEAP again. The program used provided as output for each criterion the two best sets of variables for each equation size from 1 to the number of input variables, and the coefficients of the best subset. The procedure was begun by dividing the 60 variables into 3 groups of 20 and using RLEAP on each group. The three groups are shown in Table IX. Both $\bar{R}^2$ and Cp statistics were used to search for best subsets. It was noted that the Cp statistic selected an equation of fewer variables than did the adjusted R^2 criterion in most cases. In no case did the Cp criterion select an equation of more variables than did the adjusted R^2 criterion. Because the purpose of this procedure is to find as small a subset as possible, the Cp criterion was used to select the best equation to be passed on to the next step, but the best adjusted R^2 equation was also observed because that is a more easily interpreted statistic. The subsets of variables selected from the first three groups are shown in Table X along with the Cp for each subset. The three subsets selected were then combined into one set of 27 variables and again run through

Table IX

Three Groups of Variables in First Run of RLEAP

Group 1	Group 2	Group 3
UP	NF· V	DIG· %SS
V	NF· W	DIG· PD
W	NF· CC	AN· UP
CC	NF· UP	AN· V
%SS	NB· V	EM· UP
NF	NB· W	EM· V
NB	NB· CC	EM· W
NC	NB· PD	EM· CC
SF	NC· UP	EM· PD
SB	NC· V	XMTR· UP
CF	NC· W	XMTR· V
CB	SF· V	XMTR· W
DIG	SF· W	XMTR· CC
AN	SF· CC	XMTR· %SS
EM	SF· %SS	XMTR· PD
PS	DIG· UP	BF· UP
XMTR	DIG· V	BF· V
BF	DIG· W	BF· W
NF· UP	DIG· CC	BF· %SS

Table X

Subset of Variables of Three Groups Selected by RLEAP

Group 1	Group 2	Group 3
UP	NF· W	DIG· %SS
W	NF· CC	DIG· PD
CC	NB· V	AN· W
NB	NB· W	EM· UP
SF	NC· V	EM· V
SB	NC· W	EM· W
DIG	SF· V	EM· PD
XMTR	SF· W	
	DIG· W	
Cp = 5.909	Cp = 2.958	Cp = 3.147

RLEAP to pick a smaller subset. The subset selected by this run is shown in Table XI below.

These 13 variables were then used as a core to add all of the remaining variables into three groups to give all variables a second chance to enter the equation. The new groups for this third run were formed by first including the 13 variables from the second run and then adding 15 or 16 of the 47 remaining variables until all are included in one group. Three groups of 28, 29, and 29 variables

Table XI

Subset Selected from 27 Variables Combined from Run 1 Based on Cp Criterion

UP	NF·CC	EM·
W	DIG·%SS	BF·W
SF	DIG·PD	BF·%SS
SB	AN·W	
DIG	EM·W	
$\bar{R}^2$ .83	Cp = 12.638	

resulted. When RLEAP was used on these three groups, the subsets selected by the Cp criterion shown in Table XII resulted. Asterisks indicate that the variable is one of the 13 identified by the previous run and input into all three groups. If the number of asterisks in a column is near 13, it indicates that the additional variables had little to offer in improving the model and that the model input from the previous run was relatively stable. If there are few asterisks, the additional variables had a lot to offer the model.

Until this point, the Cp statistic has been used as the criterion for choosing the best subset because the objective has been to eliminate a large number of variables quickly. But the subsets from Run 3 are small enough and the goodness-of-fit good enough for both $\bar{R}^2$ and Cp criterion that both deserved further analysis. Table XIII

Table XII

Three Subsets Resulting from the Third Run
Based on Cp Statistic

Group 1	Group 2	Group 3
W *	UP *	UP *
CC	W *	W *
%SS	SF *	SF *
NB	SB *	SB *
SF *	NF· W	NF· CC *
SB *	NF· CC *	DIG· CC
NF· UP	NF· PD	DIG· PD *
NF· CC *	NC· UP	AN· UP
DIG· PD *	NC· V	EM· W *
BF· W *	DIG· UP	EM· PD *
BF· %SS *	DIG· PD *	BF· W *
	AN· W *	BF· %SS *
	EM· W *	
	EM· PD *	
	BF· W *	
	BF· %SS *	
Cp = 6.802 $\bar{R}^2$ = .8512	Cp = 12.289 $\bar{R}^2$ = .88	Cp = 4.193 $\bar{R}^2$ = .8436

*Indicates the variable remains from the 13 from run 2.

Table XIII

Three Subsets Resulting from the Third Run
Based on $\overline{R}^2$ Criterion

Group 1		Group 2		Group 3	
UP	*	UP	*	UP	*
W	*	W	*	W	*
CC		SF	*	SF	*
%SS		SB	*	SB	*
NB		DIG	*	DIG	*
SF	*	NF*W		NF*CC	*
SB	*	NF*CC	*	DIG*V	
DIG	*	NF*PD		DIG*W	
EM		NC*UP		DIG*%SS	*
XMTR		NC*V		AN*W	*
NF*UP		SF*CC		EM*W	*
NF*CC	*	DIG*UP		EM*PD	*
DIG*%SS	*	DIG*%SS	*	BF*W	*
DIG*PD	*	DIG*PD	*	BF*%SS	*
EM*W	*	AN*W	*		
BF*W	*	EM*W	*		
BF*%SS	*	EM*PD	*		
		BF*W	*		
		BF*%SS	*		
Cp = 9.632 $\overline{R}^2$ = .8651		Cp = 12.52 $\overline{R}^2$ = .8954		Cp = 6.0 $\overline{R}^2$ = .8466	

*Indicates the variable remains from the 13 from Run 2.

contains the three sets of variables selected in Run 3 by RLEAP using the $\bar{R}^2$ criterion. The Cp values for the six sets in Tables XII or XIII can not be compared to each other because they have not been standardized as explained in Chapter I. But this will not cause a problem because the $\bar{R}^2$ has a standard basis for all cases, and because the sets are going to be run through RLEAP again in various combinations as indicated in Figure 2. On the upper branch after Run 2 are the sets selected in Run 3 by the Cp statistic. Group 2 and Group 3 on this branch will be combined in Run 4 because they are similar. The set resulting from Run 4 will then be combined in Run 5 with the set in Group 1. By that point, the model should be stabilized if it is going to. The same scheme will be used on the $\bar{R}^2$ branch of Run 3.

Table XIV shows the variables selected in each group and each run after the branch to Run 3. Figure 2 shows the branching with the results at each node in terms of number of variables, Cp value, and $\bar{R}^2$. It can be seen in Table XIV that in both the C_p and $\bar{R}^2$ branches that Group 2 dominates the model. In the C_p branch the sets selected in Run 4 combining Groups 2 and 3 and Run 5 combining Group 1 and results from Run 4 are identical to Group 2. Apparently then Group 2 is the optimum subset of the variables in Run 3 from a C_p standpoint.

On the $\bar{R}^2$ branch, Groups 2 and 3 combined in Run 4 to form a subset similar to Group 2. When Group 1 was combined in Run 5 with the set resulting from Run 4, the set selected was identical to the

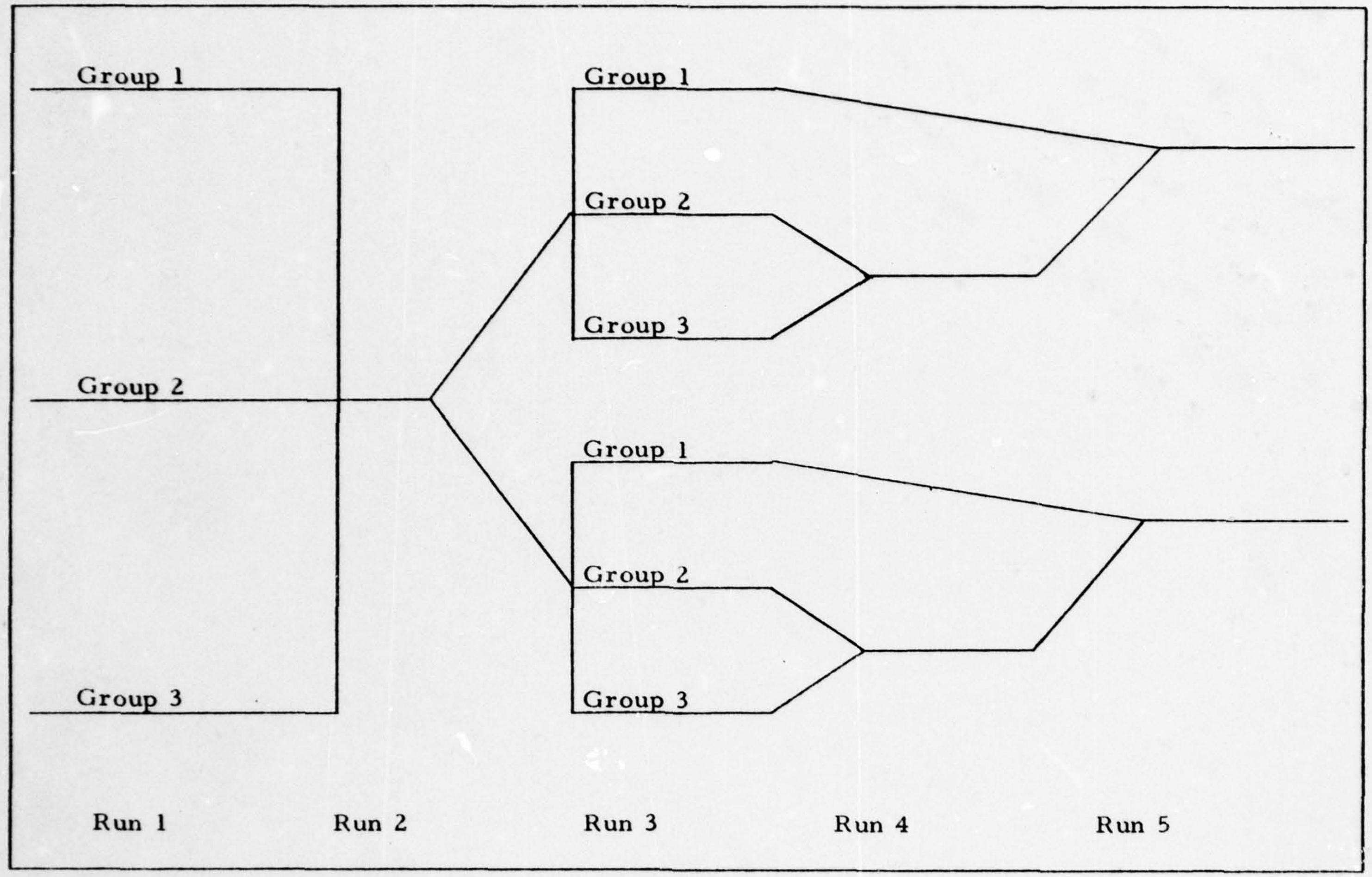


Figure 2. Sequence of Leaps and Bounds Runs

Table XIV

Variables Selected in Runs 3, 4, and 5

Variable	C_p Criterion					$\bar{R}^2$ Criterion				
	Run 3			Run 4	Run 5	Run 3			Run 4	Run 5
	Grp 1	Grp 2	Grp 3	Grp 2, 3	Grp 1, 2&3	Grp 1	Grp 2	Grp 3	Grp 2, 3	Grp 1, 2&3
UP		x	x	x	x	x	x	x	x	x
W	x	x	x	x	x	x	x	x	x	x
CC	x					x				
%SS	x					x				
NB	x					x				
SF	x	x	x	x	x	x	x	x	x	x
SB	x	x	x	x	x	x	x	x	x	x
DIG						x	x	x	x	x
EM						x				
XMTR						x				
NF*UP	x					x				
NF*W		x		x	x		x		x	x
NF*CC	x	x	x	x	x	x	x	x	x	x
NF*PD		x		x	x		x		x	x
NC*UP		x		x	x		x		x	x
NC*V		x		x	x		x		x	x
SF*CC							x		x	x
DIG*UP		x		x	x		x		x	x
DIG*V								x	x	x
DIG*W								x	x	x
DIG*CC		x								
DIG*%SS						x	x	x		
DIG*PD	x	x	x	x	x	x	x		x	x
AN*UP			x							
AN*W		x		x	x		x	x	x	x
EM*W		x	x	x	x	x	x	x	x	x
EM*PD		x	x	x	x		x	x	x	x
BF*W	x	x	x	x	x	x	x	x	x	x
BF*%SS	x	x	x	x	x	x	x	x	x	x

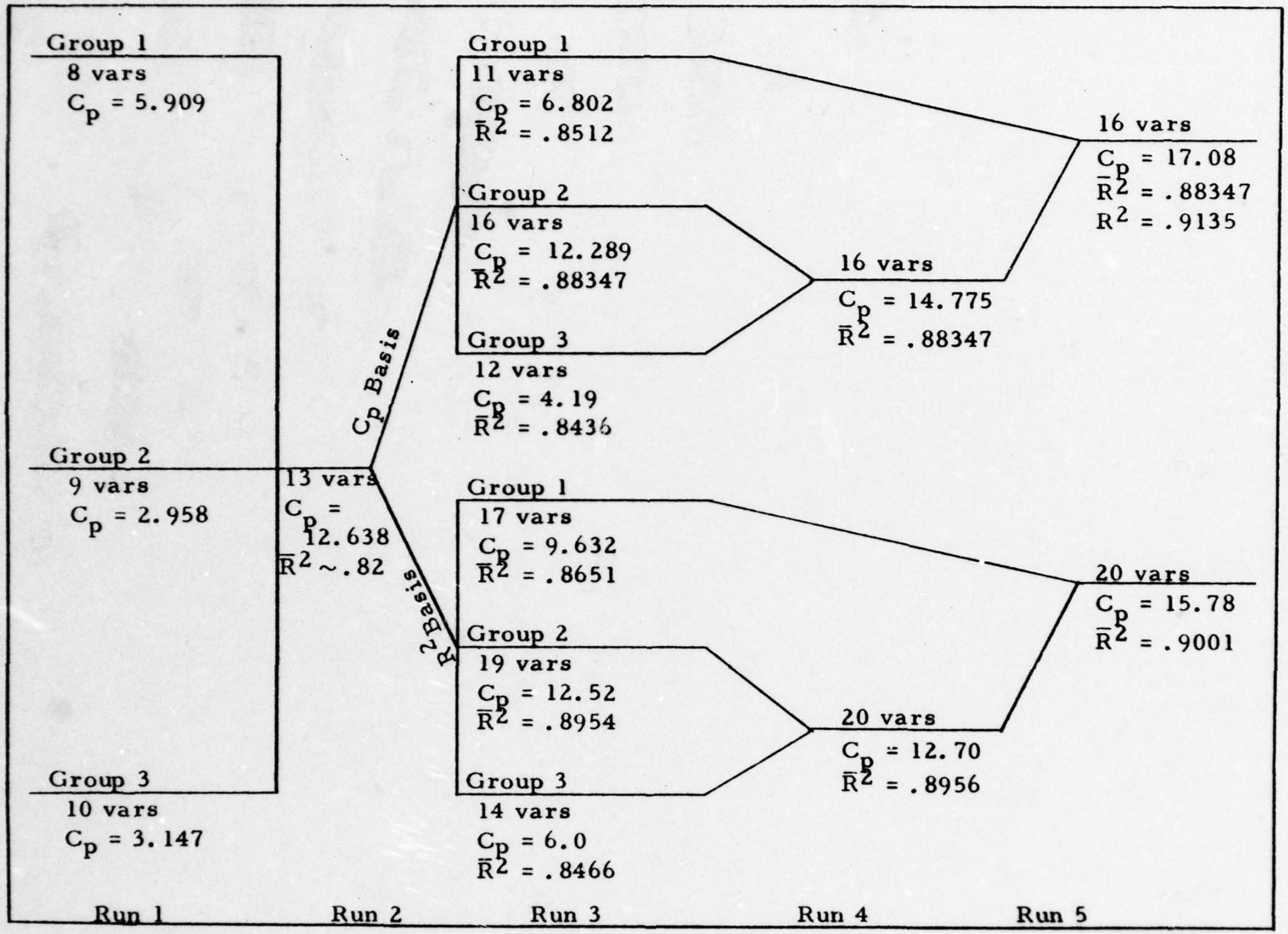


Figure 3. Results of Leaps and Bounds Runs

result from Run 4. Then the set from Run 4 must the optimal subset of the variables in Run 3 from a $\bar{R}^2$ standpoint.

The two equations found by the Leaps and Bounds method are described in Table XV, where the coefficients and partial F-values are given. It can be seen that the set generated using the C_p criterion is a subset of that generated by the $\bar{R}^2$ method. The decision as to which equation is better will be deferred until Chapter V, when the SPSS generated equation will also be considered.

One thing to keep in mind is, as Aitken said (Ref 1:226), "There may be no one "best" equation, only a most appropriate one of several adequate equations."

Table XV

Equations Selected Using the All-Possible Regressions Method

$$\ln(LSC/OH) = \alpha_0 + \sum_i \alpha_i D_i + \sum_j \beta_{j0} \ln x_j + \sum_j \sum_i \beta_{ji} D_i \ln x_j$$

Variable	C_p Criterion $R^2 = 0.9135$ $\bar{R}^2 = 0.88347$ $F = 31.21$		$\bar{R}^2$ Criterion $R^2 = 0.9323$ $\bar{R}^2 = 0.9001$ $F = 29.25$	
	Coefficient	Partial F	Coefficient	Partial F
UP	0.245908	8.78	0.313871	14.52
W	0.384075	7.75	0.350494	6.86
SF	-1.061926	12.78	-2.878942	14.29
SB	-1.822390	30.26	-2.195891	39.06
DIG			4.381530	4.88
NF*W	-0.431742	31.61	-0.343076	2.10
NF*CC	-0.466254	13.70	-0.470354	15.84
NF*PD	0.738901	16.62	0.672722	14.59
NC*UP	0.285409	5.13	0.254284	4.04
NC*V	-0.334677	4.93	-0.292486	3.92
SF*CC			0.293229	6.30
DIG*UP	-0.584870	12.86	-0.950128	11.70
DIG*V			-0.971576	2.25
DIG*W			2.676919	4.93
DIG*PD	1.081951	15.97	0.553008	2.59
AN*W	0.309271	16.60	0.239272	9.98
EM*W	0.698175	13.89	0.705835	13.47
EM*PD	-0.555855	21.58	-0.545678	20.61
BF*W	0.866668	28.67	0.828916	27.04
BF*%SS	-0.701034	37.03	-0.706378	38.19
Constant	-3.855040	53.44	-4.091618	64.16

All other coefficients are zero.

V. Conclusions

Results

Three equations have been found by stepwise regression and all possible regressions methods. They differ somewhat but have some terms in common and have similar performance characteristics. To aid in making a choice between the three, the variables contained in each and the pertinent statistics are summarized in Table XVI. The stepwise result has the higher correlation coefficients and overall F value, but it also has the most variables. Recall from Table VII that when the stepwise equation had 20 variables, as does the Leaps and Bounds $\bar{R}^2$ equation, that the R^2 was only 0.913 and $\bar{R}^2$ only 0.872. These values are somewhat lower than those for the Leaps and Bounds equations. This makes it appear that the Leaps and Bounds algorithm yields more efficient equations.

Validation

Another way to distinguish between the two equations would be to use a validation procedure on all of them. There are six data points on LRUs which were not used in finding the models, which the equations will be tried against to see how well they predict. The data points are shown in Table XVII.

Table XVI

Comparison of Three Regressions

Variable	SPSS	L&B Cp	L&B $\overline{R}^2$
UP	0.4027	0.2459	0.3139
W	0.0845	0.3841	0.3505
%SS	0.4124		
NB	11.3207		
SF	-1.1354	-1.0619	-2.8789
SB	-1.4578	-1.8224	-2.1959
DIG	3.7105		4.3815
EM	-2.9510		
PS	-0.0927		
NF·UP	0.3220		
NF·W		-0.4317	-0.3431
NF·CC	-0.5681	-0.4662	-0.4704
NF·PD		0.7389	0.6727
NB·UP	-0.7298		
NB·V	-1.8032		
NB·W	2.5068		
NC·UP		0.2854	0.2543
NC·V		-0.3347	-0.2925
SF·CC			0.2932
DIG·UP		-0.5849	-0.9501
DIG·V	-1.9960		-0.9718
DIG·W	3.0350		2.6769
DIG·PD		1.0820	0.5530
AN·UP	-0.2721		
AN·W	0.7582	0.3093	0.2893
EM·V	0.4224		
EM·W		0.6982	0.7058
EM·PD		-0.5558	-0.5457
XMTR·CC	0.2948		
XMTR·%SS	-0.4561		
BF·W	0.6979	0.8667	0.8289
BF·%SS	-0.6427	-0.7010	-0.7064
Constant	-5.3154	-3.8550	-4.0916
	$R^2 = 0.95212$ $\overline{R}^2 = 0.9239$ $F = 33.72$	$R^2 = 0.9135$ $\overline{R}^2 = 0.8835$ $F = 31.21$	$R^2 = 0.9323$ $\overline{R}^2 = 0.9001$ $F = 29.25$

Table XVII

Data Points Used for Validation

	F4E	F4E	F4E	F4E	F111	F111
UP	8046	3398	9831	9910	5514	6650
V	585.2	323.7	562.7	776.5	1025.2	738.7
W	12.5	9.3	11.8	40.8	43.3	19.0
CC	878	58	209	73	1379	900
%SS	%87	%100	%77	%69	%78	%100
PD	58	75	24	1800	311.6	253.8
DIG	0	0	0	0	0	0
ANALOG	1	0	1	1	1	1
EM	1	0	1	1	0	0
PS	0	1	0	0	0	0
XMTR	0	0	0	0	1	1
BF	0	0	1	1	0	0
TYPE USE	SF	SF	SF	SF	NF	SF

While we could get point estimates plugging the data points into the regression equations we have generated, we would have no feel for the variability of our estimates. Prediction intervals could be generated that would give a level of confidence about the estimates. None of the packages mentioned up to the point provide the data needed

to calculate prediction intervals readily. But another simple statistics package, OMNITAB, written by the National Bureau of Standards, does. The formula used to calculate prediction intervals in this case is

$$PI = \hat{\underline{y}} + t_{\frac{a}{2}, N-K-1} \left(\sqrt{(SDPV)^2 + S^2} \right) \qquad (57)$$

where $\hat{\underline{y}}$

y is the point estimate

SDPV is the Standard Deviation of the Predicted Value from OMNITAB and

S is the residual standard deviation.

All three of these values can be read directly from the OMNITAB output making the determination of prediction intervals quite easy. A listing of the OMNITAB output can be found in the Appendix in Figure 10. The OMNITAB package (Ref 18) will not perform any selection of variables, but once a subset has been decided on, it will provide useful information such as plots and the standard deviation measures needed to calculate the prediction intervals.

Table XVIII contains the prediction intervals of LSC/OH in terms of dollars/operating hour generated by each of the three equations for a 90% confidence interval.

Table XVIII

Generated by the Stepwise and the Leaps and Bounds Equations

	1	2	3	4	5	6
Actual	.589	.111	.513	.403	2.026	1.093
L&B Lower C_p	.0868	.0478	.0540	.0816	.4033	.1797
L&B C_p	.2350	.1273	.1504	.2432	1.1359	.4912
L&B Upper C_p	.6360	.3392	.4187	.7248	3.1992	1.3427
L&B Lower $\bar{R}$	.1253	.0328	.0477	.0643	.3786	.2376
L&B $\bar{R}$	.3314	.0866	.1252	.1916	1.0043	.6209
L&B Upper $\bar{R}$	.8765	.2284	.3288	.5709	2.664	1.6222
SPSS Lower	.0906	.1224	.0172	.2185	.2730	.1523
SPSS	.2089	.3068	.0431	.5218	.6658	.3619
SPSS Upper	.4818	.7689	.1079	1.247	1.6242	.8598
Westinghouse	.2952	.0909	.1313	.1598	1.2879	1.0418

The most noticeable result of comparing prediction intervals is that the SPSS equation generates 90% intervals that contain only one of the six actual values, while both Leaps and Bounds generated intervals contained five of the six actual values. On this basis, either of the Leaps and Bounds equations appears to be a better predictor than the SPSS equation. A closer look is required to choose between the two Leaps and Bounds equations, though.

First, the distance of the point estimate of each from the actual value was considered. In four cases, LRUs 1, 3, 4, and 5, the equation based on the C_p criterion had the point estimate nearest the actual value. Then the width of the 90% prediction interval was compared between the C_p and R^2 based criterion. Each criterion was found to have the narrower width of the prediction interval three times.

The C_p based equation appears to be the best predictor of those considered, based on generating the smallest error from actual values and its smaller size, 16 variables versus 20 or 23. This equation came closer than the Westinghouse equation to the actual values on two of the six cases, but Westinghouse did not generate prediction intervals to compare to. Two words of caution are necessary though. First, that the confidence levels used here can not be guaranteed to be known exactly. Refer to page 28 for the discussion of α levels. Secondly, the six LRUs used for validation were all on fighter aircraft in usage category SF or NF and none were digital.

But with the information available, it appears that the personnel of the Air Force Avionics Laboratory could perform this same type of analysis using first the Chow test to pre-select variables for elimination, then applying the Leaps and Bounds method described in Chapter IV. The data collection is the most time consuming function and could still be contracted out. Computer analysis of the data as was done in this report could be done for less than $100 of computer time using

packages already existing on the CDC system at Wright-Patterson AFB.

Recommendations

At the present time, the Westinghouse Electric Corporation is in the process of enlarging the data base and performing another analysis. In progress reports to the Avionics Laboratory, they have mentioned the possibility of finding prediction intervals which could be compared to those generated for this report. In any case, the personnel of the Avionics Laboratory could use the enlarged data base to perform their own analysis as described in Chapter IV and if results are suitable, consider generating their own cost estimating relationships on future data.

All of the methods used in this report are based on minimizing the sum of squared errors. Other criterion for finding optimal subsets have been discussed in the literature in recent years but as yet not packaged for easy use. These include Mean Square Error of Prediction (Ref 2:469) and (Ref 6:46) and Average Estimated Variance (Ref 15:261). As these methods become proven, they should be considered for use on the data base, as they place more emphasis on prediction than do the criterion used in this study.

Bibliography

1. Aitken, Murray A. "Simultaneous Inference and the Choice of Variable Subsets in Multiple Regression," Technometrics, 16: 221-227 (May 1974).

2. Allen, David M. "Mean Square Error of Prediction as a Criterion for Selecting Variables," Technometrics, 13: 469-475 (August 1971).

3. Allen, David M. "The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction," Technometrics, 16:125-127 (February 1974).

4. Andrews, D. F. "A Robust Method for Multiple Linear Regression," Technometrics, 16: 523-531 (November 1974).

5. Beale, E. M. L. et al. "The Discarding of Variables in Multivariate Analysis," Biometrika, 54: 357-366 (1967).

6. Bendel, Robert B. and A. A. Afifi. "Comparison of Stopping Rules in Forward 'Stepwise' Regressions," Journal of the American Statistical Association, 72: 46-53 (March 1977).

7. Chow, Gregory C. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," Econometrica, 28: 591-605 (July 1960).

8. Daniel, Cuthbert and Fred S. Wood. Fitting Equations to Data. New York: John Wiley and Sons, Inc. (1971).

9. Diehr, George and Donald R. Hoflin. "Approximating the Distribution of the Sample R^2 in Best Subset Regressions," Technometrics, 16: 317-320 (May 1974).

10. Dixon, W. J. BMD Biomedical Computer Programs. Los Angeles: University of California Press (1971).

11. Draper, N. R. and Smith, H. Applied Regression Analysis. New York: John Wiley and Sons, Inc. (1966).

12. Fisher, Franklin M. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note," Econometrica, 38: 361-366 (March 1970).

13. Furnival, George M. and Robert W. Wilson. "Regression by Leaps and Bounds," Technometrics, 16: 499-511 (November 1974).

14. Geary, R. C. and C. E. V. Leser. "Significance Tests in Multiple Regression," The American Statistician, 22: 20-22 (February 1968).

15. Helms, Ronald W. "The Average Estimated Variance Criterion for the Selection-of-Variables Problem in General Linear Models," Technometrics, 16: 261-273 (May 1974).

16. Hocking, R. R. "Criteria for Selection of a Subset Regression: Which One Should Be Used," Technometrics, 14: 967-70 (1972).

17. Hocking, R. R. "The Analysis and Selection of Variables in Linear Regression," Biometrics, 32: 1-49 (March 1976).

18. Hogben, David et al. OMNITAB II Users Reference Manual. Washington, D.C.: National Bureau of Standards (1971).

19. Mallows, C. L. "Some Comments on C_p," Technometrics, 15: 661-675 (1973).

20. Nie, Norman H. et al. Statistical Package for the Social Sciences. New York: McGraw Hill Book Company (1975).

21. Spjotvolle, E. "Alternatives to Plotting C_p in Multiple Regression," Biometrika, 64: 1-8 (1977).

22. Turek, John P. et al. Predictive Operations and Maintenance Cost Model. For Air Force Avionics Laboratory. Westinghouse Electric Corp. (1978).

APPENDIX

```
..

      PROGRAM YSEPR(INPUT,OUTPUT,TAPE1,TAPE2)
C
C     THIS PROGRAM FORMS THE Y1 AND Y2 MATRICES USED IN
C     PROGRAM CHOW
C
C     TAPE1 IS INPUT DATA, 6VARIABLES, 13DUMMIES, AND
C     THE DEPENDENT VARIABLE IN THAT ORDER
C     FORMAT IS 4F20.14
C
C     TAPE2 IS OUTPUT
C
      DIMENSION X(63,20),Y1(63,13),Y2(63,13)
      DO 1 K=1,63
      READ(1,1000) (X(K,JJ),JJ=1,20)
    1 CONTINUE
      DO 2 JJ=1,13
      IN=IM=0
      JD=JJ+6
      DO 3 K=1,63
      IF(X(K,JD).EQ.1.)GO TO 100
      IN=IN+1
      Y1(IN,JJ)=X(K,20)
      GO TO 3
  100 IM=IM+1
      Y2(IM,JJ)=X(K,20)
    3 CONTINUE
      PRINT 2002,IN
      WRITE(2,2002)IN
      PRINT 2000,(INC,JJ,Y1(INC,JJ),INC=1,IN)
      WRITE(2,2000) (INC,JJ,Y1(INC,JJ),INC=1,IN)
      PRINT 2003,IM
      PRINT 2001,(IMC,JJ,Y2(IMC,JJ),IMC=1,IM)
      WRITE(2,2003)IM
      WRITE(2,2001)(IMC,JJ,Y2(IMC,JJ),IMC=1,IM)
    2 CONTINUE
 1000 FORMAT(4F20.14)
 2000 FORMAT(3(5X,'Y1(',2I2,')=',F11.8))
 2001 FORMAT(3(5X,'Y2(',2I2,')=',F11.8))
 2002 FORMAT(1X,'N=',I2)
 2003 FORMAT(1X,'M=',I2)
      STOP
      END
*EOR
*EOR
*EOF
..
```

Figure 4. Program YSEPR

```
      PROGRAM SUMS(INPUT,OUTPUT,TAPE1,TAPE2)
C
C     THIS PROGRAM CALCULATES ALL SUMS OF SQUARES AND
C     CROSSPRODUCTS NEEDED FOR PROGRAM CHOW
C
C     TAPE 1 IS INPUT DATA, 6VARIABLES AND 13DUMMIES AND
C     DEPENDENT VARIABLE
C     FORMAT IS 5(4F20.14)
C
C     TAPE2 IS OUTPUT
C
      DIMENSION X(20),SS0(6,6,13),SS1(6,6,13),S0(6,13),S1(6,13)
      DO 7001 JD=7,19
      JJ=JD-6
      DO 7002 J1=1,6
      S0(J1,JJ)=0.
      S1(J1,JJ)=0.
      DO 7003 J2=1,6
      SS0(J1,J2,JJ)=0.
      SS1(J1,J2,JJ)=0.
 7003 CONTINUE
 7002 CONTINUE
 7001 CONTINUE
 1001 FORMAT(4F20.14/4F20.14/4F20.14/4F20.14/4F20.14)
      PRINT 8000
 8000 FORMAT('COMPLETED MODULE1')
      DO 1 J=1,63
      READ(1,1001)(X(M),M=1,20)
      DO 2 JD=7,19
      JJ=JD-6
      IF(X(JD).EQ.1.)GO TO 4
      DO 3 J1=1,6
      S0(J1,JJ)=S0(J1,JJ)+X(J1)
      DO 3 J2=J1,6
      SS0(J1,J2,JJ)=SS0(J1,J2,JJ)+(X(J1)*X(J2))
    3 CONTINUE
      GO TO 2
    4 DO 5 J1=1,6
      S1(J1,JJ)=S1(J1,JJ)+X(J1)
      DO 5 J2=J1,6
      SS1(J1,J2,JJ)=SS1(J1,J2,JJ)+(X(J1)*X(J2))
    5 CONTINUE
    2 CONTINUE
    1 CONTINUE
      PRINT 8001
 8001 FORMAT('COMPLETED MODULE 2')
      DO 6 JD=7,19
      JJ=JD-6
      DO 7 J1=1,5
      JF=J1+1
      DO 8 J2=JF,6
```

Figure 5. Program SUMS

```
      SS0(J2,J1,JJ)=SS0(J1,J2,JJ)
      SS1(J2,J1,JJ)=SS1(J1,J2,JJ)
    8 CONTINUE
    7 CONTINUE
    6 CONTINUE
      PRINT 8002
 8002 FORMAT('MODULE 3 COMPLETED')
      DO 9 IDUM=1,2
      ITEM=IDUM-1
      DO 10 J1=1,6
      DO 11 J2=1,6
      IF(ITEM.EQ.1)GO TO 100
      PRINT 2001,(ITEM,J1,J2,JJ,SS0(J1,J2,JJ),JJ=1,13)
 2001 FORMAT(3(1X,'SS',I1,'(',2I1,I2,')',F16.10)/,3(1X,'SS',I1,'(',2I1,I
     &2,')',F16.10)/,3(1X,'SS',I1,'(',2I1,I2,')',F16.10)/,3(1X,'SS',I1,'
     &(',2I1,I2,')',F16.10)/,1X,'SS',I1,'(',2I1,I2,')',F16.10)
      WRITE(2,2001)(ITEM,J1,J2,JJ,SS0(J1,J2,JJ),JJ=1,13)
      IF(ITEM.EQ.0)GO TO 11
  100 PRINT 2001,(ITEM,J1,J2,JJ,SS1(J1,J2,JJ),JJ=1,13)
      WRITE(2,2001)(ITEM,J1,J2,JJ,SS1(J1,J2,JJ),JJ=1,13)
   11 CONTINUE
   10 CONTINUE
    9 CONTINUE
      DO 12 IDUM=1,2
      ITEM=IDUM-1
      DO 13 J1=1,6
      IF(ITEM.EQ.1)GO TO 200
      PRINT 2002,(ITEM,J1,JJ,S0(J1,JJ),JJ=1,13)
 2002 FORMAT(3(1X,'S',I1,'(',I1,I2,')',F16.10)/3(1X,'S',I1,'(',I1,I2,')'
     &,F16.10)/3(1X,'S',I1,'(',I1,I2,')',F16.10)/3(1X,'S',I1,'(',I1,I2,'
     &)',F16.10)/1X,'S',I1,'(',I1,I2,')',F16.10)
      WRITE(2,2002)(ITEM,J1,JJ,S0(J1,JJ),JJ=1,13)
      IF(ITEM.EQ.0)GO TO 13
  200 PRINT 2002,(ITEM,J1,JJ,S1(J1,JJ),JJ=1,13)
      WRITE(2,2002)(ITEM,J1,JJ,S1(J1,JJ),JJ=1,13)
   13 CONTINUE
   12 CONTINUE
      PRINT 8003
 8003 FORMAT('COMPLETED MODULE 4')
      STOP
      END
```

```
$
1
●
    PROGRAM CHOW        74/74    OPT=1                   FTN 4.6+446         12/11/78  13.11.48        PAGE      1

              PROGRAM CHOW(INPUT,OUTPUT,TAPE1,TAPE2,TAPE3,TAPE5=OUTPUT)
        C
        C     TAPE 1 IS OUTPUT FROM PROGRAM SUMS.
        C     TAPE 2 IS OUTPUT FROM PROGRAM VSEPARATOR.
        C     TAPE 3 IS DATA WITH 6 VARIABLES, 13 DUMMIES,AND 1 INDEPENDENT
        C     VARIABLES IN THAT ORDER  FORMAT(5(4F20.14/))
        C
        C     CHOW IS DIMENSIONED TO HANDLE 63 OBSERVATIONS OF 6 INDEPENDENT,
        C     13 DUMMY, AND 1 DEPENDENT VARIABLES
        C
        C         REQUIRED IMSL ROUTINES ARE:
        C                 LINV2F
        C                 VMULFF
        C
              DIMENSION Z1(63,1),W1(63,6),Z2(63,1),W2(63,6),Y1(63,1),Y2(63,1)
             &,Z1TRAN(1,63),Z2TRAN(1,63),W1TRAN(6,63),W2TRAN(6,63),
             &AMAT0(13,13),BMAT0(13,63),AMATA(14,14),BMATA(14,63),YMAT(63,1),
             &A0INV(13,13),AAINV(14,14),SS0(6,6,13),SS1(6,6,13),S0(6,13),
             &ZINTER0(13,63),ZNTERA(14,63),FINAL0(13,1),FINALA(14,1),Z1C1(63,1),
             &W1D1(63,1),Z1C0(63,1),W1D10(63,1),Z2C2(63,1),W2D2(63,1),Z2C0(63,1)
             &,W2D20(63,1),SUM1(63,1),SUM2(63,1),SUM3(63,1),SUM4(63,1),
             &SUM5(63,1),C1(1),D1(6,1),C0(1),D10(6,1),C2(1),D2(6,1),D20(6,1),X(6
             &3,19),S1(6,13),WORK(300)
        C
        C     THIS SECTION READS SUMS OF SQUARES FROM TAPE1
        C
              DO 1 I=1,6
              DO 1 J=1,6
              READ(1,50)(SS0(I,J,JJ),JJ=1,13)
            1 CONTINUE
              DO 2 I=1,6
              DO 2 J=1,6
              READ(1,50)(SS1(I,J,JJ),JJ=1,13)
            2 CONTINUE
              DO 3 I=1,6
              READ(1,60)(S0(I,JJ),JJ=1,13)
            3 CONTINUE
              DO 4 I=1,6
              READ(1,60)(S1(I,JJ),JJ=1,13)
            4 CONTINUE
           50 FORMAT(3(10X,F16.10))
           60 FORMAT(3(8X,F16.10))
         1000 FORMAT(3X,I2)
         1001 FORMAT(3(14X,F11.8))
         1010 FORMAT(1X,*JJ=*,I2,2X,*NJJ=*,I2,2X,*MJJ=*,I2)
        C
        C         READ X DATA FROM TAPE 3
        C
```

Figure 6. Program CHOW

```
      DO 2000 K=1,63
      READ(3,3000)(X(K,I),I=1,19)
 2000 CONTINUE
 3000 FORMAT(4F20.14/4F20.14/4F20.14/4F20.14/3F20.14)
C
C        JD IS THE DUMMY VARIABLE COUNTER
C
      DO 9999 JD=7,19
      JJ=JD-6
```


```
C
C        READ SEPARATEDY FROM TAPE 2 INTO Y1 AND Y2
C
      READ(2,1000)NJJ
      READ(2,1001)(Y1(INC,1),INC=1,NJJ)
      READ(2,1000)MJJ
      READ(2,1001)(Y2(INC,1),INC=1,MJJ)
C
C        JV IS THE INDEPENDENT VARIABLE INDICATOR
C
      DO 9998 JV=1,7
      JVLESS=JV-1
      IF(JV.GT.1)GO TO 3999
C
C        FILL AMAT0, AMATA TO TEST ALPHA(JD) TERM
C
      AMAT0(1,1)=63
      AMATA(1,1)=NJJ
      AMATA(2,2)=MJJ
      DO 3001 I=1,6
      AMAT0(I+1,1)=S0(I,JJ)
      AMAT0(1,I+1)=S0(I,JJ)
      AMATA(I+2,1)=S0(I,JJ)
      AMATA(1,I+2)=S0(I,JJ)
      AMAT0(I+7,1)=S1(I,JJ)
      AMAT0(1,I+7)=S1(I,JJ)
      AMATA(I+8,2)=S1(I,JJ)
      AMATA(2,I+8)=S1(I,JJ)
 3001 CONTINUE
      DO 3002 I=1,6
      DO 3002 J=1,6
      AMAT0(I+1,J+1)=SS0(I,J,JJ)
      AMATA(I+2,J+2)=SS0(I,J,JJ)
      AMAT0(I+7,J+7)=SS1(I,J,JJ)
      AMATA(I+8,J+8)=SS1(I,J,JJ)
 3002 CONTINUE
C
C        FILL Z1, Z2, U1, AND U2 TO TEST ALPHA(JD) TERM
C
      DO 3003 I=1,NJJ
      Z1(I,1)=1.0
 3003 CONTINUE
      DO 3004 I=1,MJJ
      Z2(I,1)=1.0
 3004 CONTINUE
      IN=IM=0
      DO 3100 INC=1,63
      IF(X(INC,JD).EQ.1.0)GO TO 3150
      IN=IN+1
```

```
      DO 3110 J=1,6
      U1(IM,J)=X(INC,J)
 3110 CONTINUE
      GO TO 3100
 3150 IM=IM+1
      DO 3160 J=1,6
      U2(IM,J)=X(INC,J)
 3160 CONTINUE
```



```
C
C        EXIT TO EVALUATION SECTION
C
 3100 CONTINUE
      GO TO 8104
 3999 IF(JV.GT.2)GO TO 4071
C
C        FILL AMAT0,AMATA TO TEST BETA(1,JD) TERM
C
      DO 4001 J=9,14
      AMATA(J,1)=0.0
      AMATA(1,J)=0.0
 4001 CONTINUE
      AMAT0(1,1)=SS0(1,1,JJ)+SS1(1,1,JJ)
      AMATA(1,1)=SS0(1,1,JJ)
      AMATA(2,2)=SS1(1,1,JJ)
      AMAT0(2,1)=S0(1,JJ)
      AMATA(3,1)=S0(1,JJ)
      DO 4000 I=3,7
      AMAT0(I,1)=SS0(1,I-1,JJ)
      AMATA(I+1,1)=SS0(1,I-1,JJ)
 4000 CONTINUE
      AMAT0(8,1)=S1(1,JJ)
      AMATA(9,2)=S1(1,JJ)
      DO 4010 I=9,13
      AMAT0(I,1)=SS1(1,I-7,JJ)
      AMATA(I+1,2)=SS1(1,I-7,JJ)
 4010 CONTINUE
      AMAT0(2,2)=MJJ
      AMATA(3,3)=MJJ
      DO 4020 I=3,7
      AMAT0(I,2)=S0(I-1,JJ)
      AMATA(I+1,3)=S0(I-1,JJ)
 4020 CONTINUE
      DO 4030 J=3,7
      DO 4030 I=J,7
      AMAT0(I,J)=SS0(I-1,J-1,JJ)
      AMATA(I+1,J+1)=SS0(I-1,J-1,JJ)
 4030 CONTINUE
      AMAT0(8,8)=MJJ
      AMATA(9,9)=MJJ
      DO 4040 I=9,13
      AMAT0(I,8)=S1(I-7,JJ)
      AMATA(I+1,9)=S1(I-7,JJ)
 4040 CONTINUE
      DO 4050 J=9,13
      DO 4050 I=J,13
      AMAT0(I,J)=SS1(I-7,J-7,JJ)
      AMATA(I+1,J+1)=SS1(I-7,J-7,JJ)
 4050 CONTINUE
```

```
      DO 4060 J=1,13
      DO 4060 I=1,13
      AMATO(J,I)=AMATO(I,J)
 4060 CONTINUE
      DO 4070 J=1,14
      DO 4070 I=1,14
      AMATA(J,I)=AMATA(I,J)
 4070 CONTINUE
C
C        EXIT TO FILL Z1, Z2, W1, WZ
C
      GO TO 3050
C
C        FILL AMATO, AMATA TO TEST BETA(JV,JJ),JV=2 TO 6
C
 4071 AMATO(1,1)=SS0(JVLESS,JVLESS,JJ)+SS1(JVLESS,JVLESS,JJ)
      AMATA(1,1)=SS0(JVLESS,JVLESS,JJ)
      AMATA(2,2)=SS1(JVLESS,JVLESS,JJ)
      AMATO(2,1)=S0(JVLESS,JJ)
      AMATA(3,1)=S0(JVLESS,JJ)
      DO 3010 I=1,5
      IF(I.GE.JVLESS)GO TO 3005
      AMATO(I+2,1)=SS0(I,JVLESS,JJ)
      AMATA(I+3,1)=SS0(I,JVLESS,JJ)
      GO TO 3010
 3005 AMATO(I+2,1)=SS0(I+1,JVLESS,JJ)
      AMATA(I+3,1)=SS0(I+1,JVLESS,JJ)
 3010 CONTINUE
      AMATO(8,1)=S1(JVLESS,JJ)
      AMATA(9,2)=S1(JVLESS,JJ)
      DO 3020 I=1,5
      IF(I.GE.JVLESS)GO TO 3015
      AMATO(I+8,1)=SS1(I,JVLESS,JJ)
      AMATA(I+9,2)=SS1(I,JVLESS,JJ)
      GO TO 3020
 3015 AMATO(I+8,1)=SS1(I+1,JVLESS,JJ)
      AMATA(I+9,2)=SS1(I+1,JVLESS,JJ)
 3020 CONTINUE
      DO 3021 I=1,12
      AMATO(1,I+1)=AMATO(I+1,1)
 3021 CONTINUE
      DO 3022 I=1,6
      AMATA(1,I+2)=AMATA(I+2,1)
 3022 CONTINUE
      DO 3023 I=1,6
      AMATA(2,I+8)=AMATA(I+8,2)
 3023 CONTINUE
      AMATO(2,JVLESS+1)=S0(JVLESS-1,JJ)
      AMATA(3,JVLESS+2)=S0(JVLESS-1,JJ)
      DO 3030 I=1,5
      IF(I.GE.JVLESS)GO TO 3025
      AMATO(I+2,JVLESS+1)=SS0(I,JVLESS-1,JJ)
      AMATA(I+3,JVLESS+2)=SS0(I,JVLESS-1,JJ)
      GO TO 3030
 3025 AMATO(I+2,JVLESS+1)=SS0(I+1,JVLESS-1,JJ)
      AMATA(I+3,JVLESS+2)=SS0(I+1,JVLESS-1,JJ)
 3030 CONTINUE
      DO 3031 I=1,6
```

```
      AMATO(JVLESS+1,I+1)=AMATO(I+1,JVLESS+1)
 3031 CONTINUE
      DO 3032 I=1,7
      AMATA(JVLESS+2,I+2)=AMATA(I+2,JVLESS+2)
 3032 CONTINUE
      AMATO(8,JVLESS+7)=S1(JVLESS-1,JJ)
```



```
      AMATA(9,JVLESS+8)=S1(JVLESS-1,JJ)
      DO 3040 I=1,5
      IF(I.GE.JVLESS)GO TO 3035
      AMATO(I+8,JVLESS+7)=SS1(I,JVLESS-1,JJ)
      AMATA(I+9,JVLESS+8)=SS1(I,JVLESS-1,JJ)
      GO TO 3040
 3035 AMATO(I+8,JVLESS+7)=SS1(I+1,JVLESS-1,JJ)
      AMATA(I+9,JVLESS+8)=SS1(I+1,JVLESS-1,JJ)
 3040 CONTINUE
      DO 3041 I=1,6
      AMATO(JVLESS+7,I+7)=AMATO(I+7,JVLESS+7)
 3041 CONTINUE
      DO 3042 I=1,6
      AMATA(JVLESS+8,I+8)=AMATA(I+8,JVLESS+8)
 3042 CONTINUE
C
C        FILL Z1, Z2, U1, U2 TO TEST BETA TERMS
C
 3050 IN=0
      IM=0
      DO 9700 INC=1,63
      IF(X(INC,JD).EQ.1.0)GO TO 9500
      IN=IN+1
      DO 9100 J=1,6
      IF(J.EQ.JVLESS)GO TO 9000
      IF(J.GT.JVLESS)GO TO 9050
      U1(IN,J+1)=X(INC,J)
      GO TO 9100
 9000 Z1(IN,1)=X(INC,JVLESS)
      GO TO 9100
 9050 U1(IN,J)=X(INC,J)
 9100 CONTINUE
      GO TO 9700
 9500 IM=IM+1
      DO 9670 J=1,6
      IF(J.EQ.JVLESS)GO TO 9600
      IF(J.GT.JVLESS)GO TO 9650
      U2(IM,J+1)=X(INC,J)
      GO TO 9670
 9600 Z2(IM,1)=X(INC,JVLESS)
      GO TO 9670
 9650 U2(IM,J)=X(INC,J)
 9670 CONTINUE
 9700 CONTINUE
      DO 9800 INC=1,IN
      U1(INC,1)=1.0
 9800 CONTINUE
C
C        Z1, Z2, U1, U2 ARE FILLED
C
      DO 9850 INC=1,IM
      U2(INC,1)=1.0
```

```
 9850 CONTINUE
      DO 8001 I=1,13
      DO 8001 J=1,63
      BMAT0(I,J)=0.0
 8001 CONTINUE
PROGRAM CHOU        74/74   OPT=1                         FTN 4.6+446        12/11/78  13.11.48        PAGE    6

      DO 8004 I=1,14
      DO 8004 J=1,63
      BMATA(I,J)=0.0
 8004 CONTINUE
 8104 CALL TRANS(Z1,MJJ,1,Z1TRAN,63,1)
      CALL TRANS(Z2,MJJ,1,Z2TRAN,63,1)
      CALL TRANS(U1,MJJ,6,U1TRAN,63,6)
      CALL TRANS(U2,MJJ,6,U2TRAN,63,6)
      IDGT=3
      CALL LINV2F(AMAT0,13,13,A0INV,IDGT,WORK,IER)
C
C        FILL BMAT0
C
      CALL MOVE(Z1TRAN,1,MJJ,BMAT0,1,1,1,63,13,63)
      CALL MOVE(U1TRAN,6,MJJ,BMAT0,2,1,6,63,13,63)
      CALL MOVE(Z2TRAN,1,MJJ,BMAT0,1,MJJ+1,1,63,13,63)
      CALL MOVE(U2TRAN,6,MJJ,BMAT0,8,MJJ+1,6,63,13,63)
      CALL MOVE(Y1,MJJ,1,YMAT,1,1,63,1,63,1)
      CALL MOVE(Y2,MJJ,1,YMAT,MJJ+1,1,63,1,63,1)
C
C        CALCULATE VECTOR OF COEFFICIENTS FOR NULL HYPOTHESIS
C
      CALL VMULFF(A0INV,BMAT0,13,13,63,13,13,ZINTER0,13,IER)
      CALL VMULFF(ZINTER0,YMAT,13,63,1,13,63,FINAL0,13,IER)
      IDGT=3
      CALL LINV2F(AMATA,14,14,AAINV,IDGT,WORK,IER)
C
C        FILL BMATA
C
      CALL MOVE(Z1TRAN,1,MJJ,BMATA,1,1,1,63,14,63)
      CALL MOVE(U1TRAN,6,MJJ,BMATA,3,1,6,63,14,63)
      CALL MOVE(Z2TRAN,1,MJJ,BMATA,2,MJJ+1,1,63,14,63)
      CALL MOVE(U2TRAN,6,MJJ,BMAYA,9,MJJ+1,6,63,14,63)
C
C        CALCULATE VECTOR OF COEFFICIENTS FOR ALTEPNATE HYPOTHESIS
C
      CALL VMULFF(AAINV,BMATA,14,14,63,14,14,ZNTERA,14,IER)
      CALL VMULFF(ZNTERA,YMAT,14,63,1,14,63,FINALA,14,IER)
      C0(1)=FINAL0(1,1)
      DO 5000 I=1,6
      D10(I,1)=FINAL0(I+1,1)
 5000 CONTINUE
      DO 5001 I=1,6
      D20(I,1)=FINAL0(I+7,1)
 5001 CONTINUE
      C1(1)=FINALA(1,1)
      C2(1)=FINALA(2,1)
      DO 5002 I=1,6
      D1(I,1)=FINALA(I+2,1)
 5002 CONTINUE
      DO 5003 I=1,6
      D2(I,1)=FINALA(I+8,1)
 5003 CONTINUE
```

```
C
C          CALCULATE F VALUES
C
      CALL VMULFF(Z1,C1,MJJ,1,1,63,1,Z1C1,63,IER)
1          PROGRAM CHOW       74/74    OPT=1                    FTN 4.6+446         12/11/78  13.11.48        PAGE     7
      CALL VMULFF(U1,D1,MJJ,6,1,63,6,U1D1,63,IER)
      CALL VMULFF(Z1,C0,MJJ,1,1,63,1,Z1C0,63,IER)
      CALL VMULFF(U1,D10,MJJ,6,1,63,6,U1D10,63,IER)
      CALL ADD(Z1C1,U1D1,MJJ,1,SUM1,63,1)
      CALL SUBTR(SUM1,Z1C0,MJJ,1,SUM1,63,1)
      CALL SUBTR(SUM1,U1D10,MJJ,1,SUM1,63,1)
      CALL VMULFF(Z2,C2,MJJ,1,1,63,1,Z2C2,63,IER)
      CALL VMULFF(U2,D2,MJJ,6,1,63,6,Z2C2,63,IER)
      CALL VMULFF(Z2,C0,MJJ,1,1,63,1,Z2C0,63,IER)
      CALL VMULFF(U2,D20,MJJ,6,1,63,6,U2D20,63,IER)
      CALL ADD(Z2C2,U2D2,MJJ,1,SUM2,63,1)
      CALL SUBTR(SUM2,Z2C0,MJJ,1,SUM2,63,1)
      CALL SUBTR(SUM2,U2D20,MJJ,1,SUM2,63,1)
      CALL SUBTR(Y1,Z1C1,MJJ,1,SUM3,63,1)
      CALL SUBTR(SUM3,U1D1,MJJ,1,SUM3,63,1)
      CALL SUBTR(Y2,Z2C2,MJJ,1,SUM4,63,1)
      CALL SUBTR(SUM4,U2D2,MJJ,1,SUM4,63,1)
      CALL SUBTR(Y2,Z2C0,MJJ,1,SUM5,63,1)
      CALL SUBTR(SUM5,U2D20,MJJ,1,SUM5,63,1)
      CALL NORM(SUM1,MJJ,SONE,63)
      CALL NORM(SUM2,MJJ,STWO,63)
      CALL NORM(SUM3,MJJ,STHREE,63)
      CALL NORM(SUM4,MJJ,SFOUR,63)
      CALL NORM(SUM5,MJJ,SFIVE,63)
      IF(MJJ.GT.7)GO TO 5010
      IF(MJJ.LT.6)GO TO 5015
C
C          P-Q .LE. MJJ .LE.P
C
      F=(SONE+SFIVE)/STHREE*(MJJ-6)/(MJJ-6.)
      WRITE(5,5005)JJ,JVLESS,MJJ,F
 5005 FORMAT(/,1X,'CURRENT INDICATOR=',I2,2X,'TESTED VARIABLE=',I1,2X,'P
     &OINTS IN SET 2=',I2,2X,'F VALUE=',F14.6)
      GO TO 9998
C
C          M > P
C
 5010 F=(SONE+STWO)/(STHREE+SFOUR)*49.
      WRITE(5,5005)JJ,JVLESS,MJJ,F
      GO TO 9998
C
C          M < P-Q
C
 5015 WRITE(5,5016)MJJ,JJ,JVLESS
 5016 FORMAT(1X,'SECOND SET CONTAINS',1X,I1,1X,'POINTS. INSUFFICIENT FOR
     &CALCULATION OF F.  INDICATOR= ',I2,'   VARIABLE= ',I1)
 9998 CONTINUE
 9999 CONTINUE
      STOP
      END

      SYMBOLIC REFERENCE MAP (R=1)
1          PROGRAM CHOW       74/74    OPT=1                    FTN 4.6+446         12/11/78  13.11.48        PAGE     8
```

```
SUBROUTINE TRANS       74/74   OPT=1                       FTN 4.6+446         12/11/78  12.33.28       PAGE     1

      C
      C
      C
            SUBROUTINE TRANS(A,NR,NC,ATRAN,NRD,NCD)
            DIMENSION A(NRD,NCD),ATRAN(NCD,NRD)
      C
      C        A IS MATRIX TO BE TRANSPOSED
      C        NR AND NC IS NUMBER OF ROWS ANDCOLUMNS IN MATRIX A
      C        NRD AND NCD IS NUMBER OF ROWS AND COLUMNS DIMENSIONED FOR A
      C        ATRAN IS THE TRANSPOSE MATRIX
      C
            DO 1 I=1,NC
            DO 1 J=1,NR
            ATRAN(I,J)=A(J,I)
          1 CONTINUE
            RETURN
            END
```

```
SUBROUTINE MOVE        74/74   OPT=1                       FTN 4.6+446         12/11/78  12.33.28       PAGE     1

      C
      C
      C
            SUBROUTINE MOVE(XOA,IOR,IOC,XNA,INR,INC,NRD,NCD,INRD,INCD)
            DIMENSION XOA(NRD,NCD),XNA(INRD,INCD)
      C        MOVE  FILL MATRIX XNA WITH SUBMATRICES XOA
      C        XOA IS THE MATRIX TO BE MOVED
      C        IOR AND IOC ARE NUMBER OF ROWS AND COLUMNS IN XOA
      C        XNA IS THE MATRIX BEING FILLED
      C        INR AND INC ARE THE ROW AND COLUMN POSITION THAT XOA(1,1)
      C        STARTS IN MATRIX XNA
      C        NRD AND NCD ARE THE NO. OF ROWS AND COLUMNS DIMENSIONED FOR XOA
      C        INRD AND INCD ARE THE ROW AND COLUMN SIZES DIMENSIONED FOR MATRIX XNA
            DO 10 I=1,IOR
            DO 10 J=1,IOC
            NR=INR+I-1
            NC=INC+J-1
            XNA(NR,NC)=XOA(I,J)
         10 CONTINUE
            RETURN
            END
```

```
SUBROUTINE ADD          74/74    OPT=1                     FTN 4.6+446         12/11/78  13.11.48        PAGE     1

      SUBROUTINE ADD(A,B,NR,NC,C,NRD,NCD)
C
C        ADDS TWO MATRICES , A AND B, TOGETHER
C        NR AND NC ARE THE NUMBER OF ROWS AND COLUMNS IN A AND B
C        NRD AND NCD ARE THE NUMBER OF ROWS AND COLUMNS DIMENSIONED FOR A AND B
C
      DIMENSION A(NRD,NCD),B(NRD,NCD),C(NRD,NCD)
      DO 1 J=1,NC
      DO 1 I=1,NR
      C(I,J)=A(I,J)+B(I,J)
    1 CONTINUE
      RETURN
      END
SUBROUTINE SUBTR        74/74    OPT=1                     FTN 4.6+446         12/11/78  13.11.48        PAGE     1

      SUBROUTINE SUBTR(A,B,NR,NC,C,NRD,NCD)
C
C        THIS SUBTRACTS MATRIX B FROM MATRIX A
C        DIMENSIONS ARE THE SAME AS IN SUBROUTINE ADD
C
      DIMENSION A(NRD,NCD),B(NRD,NCD),C(NRD,NCD)
      DO 1 J=1,NC
      DO 1 I=1,NR
      C(I,J)=A(I,J)-B(I,J)
    1 CONTINUE
      RETURN
      END
SUBROUTINE NORM         74/74    OPT=1                     FTN 4.6+446         12/11/78  13.11.48        PAGE     1

      SUBROUTINE NORM(A,NR,B,NRD)
C
C        CALCULATES THE SQUARE OF THE NORM FOR MATRIX A
C        SQUARED NORM IS PUT INTO B
C        NR IS THE NUMBER OF ROWS IN A
C        NRD IS THE NUMBER OF ROWS DIMENSIONED FOR A
      DIMENSION A(NRD,1)
      B=0.0
      DO 1 I=1,NR
      B=A(I,1)**2+B
    1 CONTINUE
      RETURN
      END
```

```
$
1EOR
CURRENT INDICATOR- 1  TESTED VARIABLE-0  POINTS IN SET 2-12  F VALUE-     46.792030
CURRENT INDICATOR- 1  TESTED VARIABLE-1  POINTS IN SET 2-12  F VALUE-     35.957190
CURRENT INDICATOR- 1  TESTED VARIABLE-2  POINTS IN SET 2-12  F VALUE-     25.792171
CURRENT INDICATOR- 1  TESTED VARIABLE-3  POINTS IN SET 2-12  F VALUE-     25.082949
CURRENT INDICATOR- 1  TESTED VARIABLE-4  POINTS IN SET 2-12  F VALUE-     25.320975
CURRENT INDICATOR- 1  TESTED VARIABLE-5  POINTS IN SET 2-12  F VALUE-      4.445489
CURRENT INDICATOR- 1  TESTED VARIABLE-6  POINTS IN SET 2-12  F VALUE-     30.243100
CURRENT INDICATOR- 2  TESTED VARIABLE-0  POINTS IN SET 2-10  F VALUE-     48.994204
CURRENT INDICATOR- 2  TESTED VARIABLE-1  POINTS IN SET 2-10  F VALUE-     31.164791
CURRENT INDICATOR- 2  TESTED VARIABLE-2  POINTS IN SET 2-10  F VALUE-     27.367543
CURRENT INDICATOR- 2  TESTED VARIABLE-3  POINTS IN SET 2-10  F VALUE-     17.833868
CURRENT INDICATOR- 2  TESTED VARIABLE-4  POINTS IN SET 2-10  F VALUE-     37.558254
CURRENT INDICATOR- 2  TESTED VARIABLE-5  POINTS IN SET 2-10  F VALUE-      2.207488
CURRENT INDICATOR- 2  TESTED VARIABLE-6  POINTS IN SET 2-10  F VALUE-     13.168126
CURRENT INDICATOR- 3  TESTED VARIABLE-0  POINTS IN SET 2-12  F VALUE-     47.187560
CURRENT INDICATOR- 3  TESTED VARIABLE-1  POINTS IN SET 2-12  F VALUE-     35.020414
CURRENT INDICATOR- 3  TESTED VARIABLE-2  POINTS IN SET 2-12  F VALUE-     20.261425
CURRENT INDICATOR- 3  TESTED VARIABLE-3  POINTS IN SET 2-12  F VALUE-     42.982841
CURRENT INDICATOR- 3  TESTED VARIABLE-4  POINTS IN SET 2-12  F VALUE-       .434402
CURRENT INDICATOR- 3  TESTED VARIABLE-5  POINTS IN SET 2-12  F VALUE-      1.663537
CURRENT INDICATOR- 3  TESTED VARIABLE-6  POINTS IN SET 2-12  F VALUE-      2.142207
CURRENT INDICATOR- 4  TESTED VARIABLE-0  POINTS IN SET 2-12  F VALUE-     25.037979
CURRENT INDICATOR- 4  TESTED VARIABLE-1  POINTS IN SET 2-12  F VALUE-      4.028575
CURRENT INDICATOR- 4  TESTED VARIABLE-2  POINTS IN SET 2-12  F VALUE-     43.242303
CURRENT INDICATOR- 4  TESTED VARIABLE-3  POINTS IN SET 2-12  F VALUE-     42.912487
CURRENT INDICATOR- 4  TESTED VARIABLE-4  POINTS IN SET 2-12  F VALUE-     35.336394
CURRENT INDICATOR- 4  TESTED VARIABLE-5  POINTS IN SET 2-12  F VALUE-     12.666487
CURRENT INDICATOR- 4  TESTED VARIABLE-6  POINTS IN SET 2-12  F VALUE-       .138342
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 0
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 1
..
```

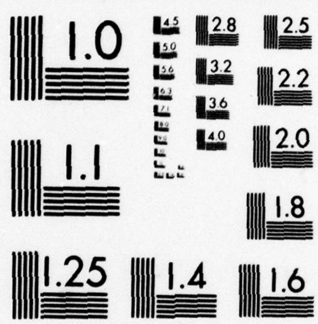
```
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 2
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 3
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 4
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 5
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  5   VARIABLE- 6
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 0
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 1
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 2
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 3
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 4
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 5
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  6   VARIABLE- 6
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 0
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 1
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 2
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 3
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 4
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 5
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR-  7   VARIABLE- 6
CURRENT INDICATOR- 8  TESTED VARIABLE-0  POINTS IN SET 2- 8  F VALUE-    49.001918

CURRENT INDICATOR- 8  TESTED VARIABLE-1  POINTS IN SET 2- 8  F VALUE-    36.238245

CURRENT INDICATOR- 8  TESTED VARIABLE-2  POINTS IN SET 2- 8  F VALUE-    36.429398

CURRENT INDICATOR- 8  TESTED VARIABLE-3  POINTS IN SET 2- 8  F VALUE-    47.933922

CURRENT INDICATOR- 8  TESTED VARIABLE-4  POINTS IN SET 2- 8  F VALUE-    44.038580

CURRENT INDICATOR- 8  TESTED VARIABLE-5  POINTS IN SET 2- 8  F VALUE-    48.997518

CURRENT INDICATOR- 8  TESTED VARIABLE-6  POINTS IN SET 2- 8  F VALUE-    35.318639

CURRENT INDICATOR- 9  TESTED VARIABLE-0  POINTS IN SET 2-51  F VALUE-    49.440065

CURRENT INDICATOR- 9  TESTED VARIABLE-1  POINTS IN SET 2-51  F VALUE-    40.300204

CURRENT INDICATOR- 9  TESTED VARIABLE-2  POINTS IN SET 2-51  F VALUE-    14.671381

CURRENT INDICATOR- 9  TESTED VARIABLE-3  POINTS IN SET 2-51  F VALUE-    45.948176

CURRENT INDICATOR- 9  TESTED VARIABLE-4  POINTS IN SET 2-51  F VALUE-      .735032

CURRENT INDICATOR- 9  TESTED VARIABLE-5  POINTS IN SET 2-51  F VALUE-     7.367655

CURRENT INDICATOR- 9  TESTED VARIABLE-6  POINTS IN SET 2-51  F VALUE-      .680680

CURRENT INDICATOR-10  TESTED VARIABLE-0  POINTS IN SET 2-26  F VALUE-    48.900573

CURRENT INDICATOR-10  TESTED VARIABLE-1  POINTS IN SET 2-26  F VALUE-    35.818197

CURRENT INDICATOR-10  TESTED VARIABLE-2  POINTS IN SET 2-26  F VALUE-    43.051565

CURRENT INDICATOR-10  TESTED VARIABLE-3  POINTS IN SET 2-26  F VALUE-    21.841772

CURRENT INDICATOR-10  TESTED VARIABLE-4  POINTS IN SET 2-26  F VALUE-    28.714985

CURRENT INDICATOR-10  TESTED VARIABLE-5  POINTS IN SET 2-26  F VALUE-     7.509605
```

AD-A064 658 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 12/1
CRITERION FOR SELECTION OF VARIABLES IN A REGRESSION ANALYSIS.(U)
DEC 78 L J PULCHER

UNCLASSIFIED AFIT/GOR/MA/78D-5 NL

2 OF 2
AD
A064658

END
DATE
FILMED
4 -79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

```
CURRENT INDICATOR-10  TESTED VARIABLE-6  POINTS IN SET 2-26  F VALUE-     19.330805
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 0
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 1
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 2
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 3
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 4
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 5
SECOND SET CONTAINS 4 POINTS. INSUFFICIENT FORCALCULATION OF F.  INDICATOR- 11    VARIABLE- 6

CURRENT INDICATOR-12  TESTED VARIABLE-0  POINTS IN SET 2-23  F VALUE-     54.983835

CURRENT INDICATOR-12  TESTED VARIABLE-1  POINTS IN SET 2-23  F VALUE-     42.912617

CURRENT INDICATOR-12  TESTED VARIABLE-2  POINTS IN SET 2-23  F VALUE-     45.636063

CURRENT INDICATOR-12  TESTED VARIABLE-3  POINTS IN SET 2-23  F VALUE-     35.171661

CURRENT INDICATOR-12  TESTED VARIABLE-4  POINTS IN SET 2-23  F VALUE-     15.781559

CURRENT INDICATOR-12  TESTED VARIABLE-5  POINTS IN SET 2-23  F VALUE-     17.990961

CURRENT INDICATOR-12  TESTED VARIABLE-6  POINTS IN SET 2-23  F VALUE-     15.147350

CURRENT INDICATOR-13  TESTED VARIABLE-0  POINTS IN SET 2-20  F VALUE-     46.647943

CURRENT INDICATOR-13  TESTED VARIABLE-1  POINTS IN SET 2-20  F VALUE-     39.790991

CURRENT INDICATOR-13  TESTED VARIABLE-2  POINTS IN SET 2-20  F VALUE-     22.917883

CURRENT INDICATOR-13  TESTED VARIABLE-3  POINTS IN SET 2-20  F VALUE-     39.312276

CURRENT INDICATOR-13  TESTED VARIABLE-4  POINTS IN SET 2-20  F VALUE-     11.509766

CURRENT INDICATOR-13  TESTED VARIABLE-5  POINTS IN SET 2-20  F VALUE-     15.344824

CURRENT INDICATOR-13  TESTED VARIABLE-6  POINTS IN SET 2-20  F VALUE-     11.583250
```

```
VOGELBACK COMPUTING CENTER                                        11/24/78     12.35.34.     PAGE    1
NORTHWESTERN UNIVERSITY

S P S S - - STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES

VERSION 7.0 -- JUNE 27 1977


          RUN NAME          SPSSSTEPWISE WITH 60 LEFT FROM SHOW
          PRINT BACK        CONTROL
          VARIABLE LIST     V1 TO V93
          INPUT FORMAT      FIXED(72(12X,3F20.14/),12X,2F20.14)

THE INPUT FORMAT PROVIDES FOR  94 VARIABLES.   93 WILL BE READ
IT PROVIDES FOR 33 RECORDS (*CARDS*) PER CASE.  A MAXIMUM OF   72 *COLUMNS* ARE USED ON A RECORD.

WARNING - A NUMERIC VARIABLE HAS A WIDTH GREATER THAN 14.  SMALL ROUNDING/TRUNCATION ERRORS MAY OCCUR.
          N OF CASES        63
          INPUT MEDIUM      DISK
          REGRESSION        METHOD=STEPWISE/
                            VARIABLES=V1 TO V42, V52 TO V96/
                            REGRESSION=V93 WITH V1 TO V23, V25 TO V29, V32 TO V34,
                            V39 TO V42, V52 TO V70, V74 TO V77, V79, V56 TO V94,V96/
                            RESIDUALS/
          OPTIONS           1,19
          STATISTICS        ALL
          READ INPUT DATA
          FINISH

       00071300 CM NEEDED FOR REGRESSION


       OPTION - 1
       IGNORE MISSING VALUE INDICATORS

       OPTION -19
       PRINT UNNORMALIZED REGRESSION COEFFICIENTS
```

Figure 7. Selected Output from SPSS

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

* M U L T I P L E   R E G R E S S I O N *

DEPENDENT VARIABLE.. V93

VARIABLE(S) ENTERED ON STEP NUMBER 24.. V23

MULTIPLE R	.96172	ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIGNIFICANCE
R SQUARE	.92491	REGRESSION	22.	144.70343	6.57743	22.39496	.000
ADJUSTED R SQUARE	.88301	RESIDUAL	40.	11.74805	.29370		
STD DEVIATION	.54194	COEFF OF VARIABILITY	144.0 PCT				

-------------------- VARIABLES IN THE EQUATION --------------------

VARIABLE	B	STD ERROR B	F	SIGNIFICANCE	BETA	ELASTICITY
V3	.42922361E-01	.32573206	.17363557E-01	.896	.0290174	-.35554
V58	-.21210898	.12234727	3.0055394	.091	-.4652592	7.97146
V11	-1.4539365	.35032200	17.224514	.000	-.2243740	.24529
V64	3.1150604	.89655375	12.072029	.001	2.2652553	-3.35225
V8	10.188694	2.85.0934	12.743777	.001	2.3125991	-.28626
V45	.22371112	.69441273E-01	10.378022	.003	.4134394	-1.45232
V63	-2.0377036	.57836834	12.412055	.001	-2.9553598	4.55119
V95	-.52481753	.11590701	20.502056	.000	-.6583567	1.79709
V27	-1.7036050	.72434635	5.7314334	.024	-2.6200504	4.55049
V15	-3.1362296	1.1363529	7.7171530	.010	-.9554153	3.49506
V17	-.77959975E-01	.39032668	.39767306E-01	.843	-.0120519	.01315
V34	.59544839	.16701779	12.329178	.001	.5535109	-1.47922
V10	-1.2684846	.33202735	14.595583	.000	-.7152527	.54202
V1	.48193212	.13326647	13.077711	.001	.4793223	-11.37060
V5	.71401351	.75420443E-01	15.532341	.000	.3545321	-2.75730
V56	-.37249904	.10336377	9.1139523	.004	-.4161542	1.20642
V28	2.2715973	.85424164	7.1932525	.011	1.7275511	-2.56627
V75	.42355858	.17697539	5.7276727	.021	.5577151	-3.05143

---------- VARIABLES NOT IN THE EQUATION ----------

VARIABLE	PARTIAL	TOLERANCE	F	SIGNIFICANCE
V2	.05743	.05293	.30043745	.587
V4	-.24762	.20609	2.5474521	.119
V6	.17565	.29350	1.2445434	.271
V7	.44348	.04677	9.5536450	.004
V9	.04227	.32496	.69325265E-01	.793
V12	-.05395	.56216	.1139-091	.779
V13	-.07850	.80756	.24164524	.626
V15	-.11157	.01611	.49425870	.486
V18	.26155	.01552	2.8536356	.099
V19	.02346	.06640	.22216704E-01	.882
V20	.50195	.04926	22.162375	.000
V21	.53763	.04456	15.355336	.000
V22	.35153	.06733	5.4986740	.024
V25	.52427	.06450	14.790043	.000
V29	-.05820	.02050	.13255217	.718
V32	.01049	.30042	.42913020E-02	.948
V33	.03282	.31857	.42301241E-01	.838
V34	.01619	.31358	.10220776E-01	.920

V26	-.67951336	.32029424	3.7441724 .159	-1.1363875 2.02322
V75	.51449273	.30733950	3.6523229 .056	.63544153 -4.29721
V14	3.5341170	1.7935543	4.4511423 .041	.75931775 -1.21273
V23	-.61276435E-01	.42303636E-01	3.6055246 .065	-.11407337 .22240
(CONSTANT)	-3.1195646	.73743092	56.014556 .000	

V39	.08249	.00938	.26980226 .606
V40	.15125	.02457	.91307010 .345
V41	.21261	.04453	1.8499675 .182
V42	.21737	.09975	1.9341959 .172
V62	-.25557	.00452	2.7322176 .106
V65	-.17711	.00534	1.2630375 .268
V66	-.16317	.00001	1.0568216 .308
V67	.04545	.01074	.80730644E-01 .775
V69	.03634	.01475	.51571752E-01 .821
V74	-.12033	.01095	.57294434 .454
V76	-.19637	.01667	1.5642515 .215
V77	-.11566	.02015	.52877765 .471
V79	.06102	.02557	.14577120 .705
V86	.33432	.01955	4.9074115 .033
V87	.22584	.02115	2.0961424 .156
V88	.27251	.02531	3.1316961 .085
V91	.17375	.03776	1.1711415 .286
V92	.02531	.02378	.25088927E-01 .875
V93	.02967	.02107	.34353229E-01 .854

SPSSSTEPWISE WITH 60 LEFT FROM CHOW   11/24/78   12.35.34.   PAGE 69

FILE NONAME (CREATION DATE = 11/24/78)

* M U L T I P L E   R E G R E S S I O N *

DEPENDENT VARIABLE.. V93

VARIABLE(S) ENTERED ON STEP NUMBER 25.. V20

MULTIPLE R	.97577	ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F	SIGNIFICANCE
R SQUARE	.95212	REGRESSION	23.	146.96044	5.47654	33.71811	0
ADJUSTED R SQUARE	.92358	RESIDUAL	35.	7.49103	.19298		
STD DEVIATION	.43927	COEFF OF VARIABILITY	116.5 PCT				

-------------------- VARIABLES IN THE EQUATION --------------------

VARIABLE	B	STD ERROR B	F	SIGNIFICANCE	BETA	ELASTICITY
V3	.865484705E-01	.26356767	.10290281	.750	.0572175	-.70232
V56	-.27214195	.59760343E-01	7.4417135	.010	-.6277063	5.03592
V11	-1.4578595	.28330689	26.474313	.000	-.2255855	.21596
V64	3.0349698	.72524203	17.512793	.000	2.2133102	-3.45093
V8	11.725694	2.3205931	23.799293	.000	2.3251700	-4.77481
V89	.29403933	.59154103E-01	25.704556	.000	.6100095	-1.81400
V63	-1.9955686	.46782655	18.202773	.000	-2.2947157	4.55271
V95	-.64273556	.97022893E-01	43.885212	0	-.3332327	2.20197
V27	-1.5332419	.58016251	9.6533769	.004	-2.7132695	4.51462
V15	-2.9509702	.96070348	9.4351956	.004	-.9219173	3.23509
V17	-.327158245E-01	.31612660	.80916172E-01	.771	-.0143566	.01364
V94	.52789478	.13712608	25.902787	.000	.6763954	-1.75032
V10	-1.1354451	.25799375	17.605597	.000	-.2822717	.57569
V1	.40270193	.10907872	13.529727	.001	.4109781	-5.50141
V5	.41240693	.67525455E-01	37.281799	0	.4759895	-3.42131
V90	-.44514529	.91453867E-01	24.958276	.000	-.1193790	1.55525
V29	2.5066232	.71966273	12.203685	.001	.8232157	-2.51209
V75	.42277740	.14312140	9.7094736	.005	.6632426	-7.05269

---------- VARIABLES NOT IN THE EQUATION ----------

VARIABLE	PARTIAL	TOLERANCE	F	SIGNIFICANCE
V2	.09325	.05295	.24532645	.623
V4	.13070	.12998	1.3575136	.252
V6	.09291	.28445	.30360545	.569
V7	-.21251	.01151	1.6043290	.187
V9	.21312	.31112	1.5781775	.187
V12	.09579	.57439	.35782394	.553
V13	-.13940	.89530	.71300463	.391
V15	-.01519	.01555	.67755407E-02	.935
V18	-.07242	.01137	.20511567	.655
V19	-.01524	.06515	.10026316E-01	.921
V21	.02662	.01037	.27345574E-01	.870
V22	.03917	.04863	.53352732E-01	.810
V26	.13751	.02816	.73241045	.397
V29	-.05156	.01999	.10265790	.753
V32	.18748	.28523	1.3442908	.247
V33	.16954	.30613	1.2307663	.265
V34	.12741	.30749	.62704432	.433
V39	.16395	.03936	.44496226	.494

SPSSSTEPWISE WITH AP LEFT FROM CHOW

V25	-.72384636	.21035237	7.0084640 .009	-1.3002062 2.71-25
V20	.73323939	.26070285	8.1070295 .007	.7537770 -5.05563
V24	3.7105270	1.3773852	7.2517520 .010	.7675742 -1.25201
V23	-.55402451	.10304578	27.139695 .000	-.7973219 1.55.47
V20	.32201546	.68431652E-01	22.142576 .000	.7439581 -1.43233
(CONSTANT)	-2.7153730	.59798024	79.012258 0	

11/24/75 12.35.34. PAGE 70

V40	.16262	.02454	1.0322008 .316
V41	.17162	.04355	1.1532221 .290
V42	.20654	.09698	1.6933024 .201
V62	-.16211	.00459	1.0255353 .313
V65	-.15032	.00529	.67845769 .355
V66	-.09157	.00001	.32135392 .574
V67	.01429	.01071	.64257903E-12 .927
V69	.05527	.01475	.11543674 .735
V74	-.05144	.00931	.14401163 .706
V76	-.05262	.01556	.10549325 .747
V77	.11310	.01791	.69233363 .407
V79	.07570	.02543	.48501445E-11 .827
V86	.00919	.01377	.34229001E-12 .954
V87	-.09018	.01623	.31154340 .580
V88	-.03956	.01940	.59371562E-11 .806
V91	-.06058	.03205	.28364358 .597
V92	-.02765	.02354	.27826096E-11 .868
V93	-.06364	.02101	.50238657E-13 .982

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

SPSSSTEPWISE WITH 60 LEFT FROM CHOW 11/24/78 12.35.34. PAGE 143

FILE NONAME (CREATION DATE = 11/24/78)

* MULTIPLE REGRESSION *

DEPENDENT VARIABLE.. V94

COEFFICIENTS AND CONFIDENCE INTERVALS.

VARIABLE	B	STD ERROR B	T	95.0 PCT CONFIDENCE INTERVAL	
V3	-2.2556830	1.3557644	-1.6596376	-5.1730446 ,	.65967866
V53	-.11395714	.45032059	-.35195930	-1.1297567 ,	.90158444
V11	-1.5127731	.65064794	-2.3044922	-2.9207199 ,	-.10483579
V64	3.3242429	1.3465698	2.4450154	.63155053 ,	6.2166362
V6	13.942642	4.2531857	3.6933209	7.1151070 ,	24.570177
V89	1.0920539	.37596253	2.8779720	.27565031 ,	1.9363695
V53	-1.0492534	.72718450	-1.4429397	-2.5069793 ,	.51037223
V95	-.22790313	.31654726	-.71922052	-.91133373 ,	.45552747
V27	-2.4966039	1.2513414	-2.3143724	-5.5749301 ,	-.21220373
V17	-3.3025033	1.1192468	-2.9549749	-5.6377491 ,	-.32125393
V1	.45759846	.25544640	1.7640311	-.15034542 ,	1.0667463
V5	.17713177	.12319653	1.3461423	-.10311937 ,	.44937285
V90	-.23964602	.22579610	-.36632193	-.65223563 ,	.35375264
V28	4.1363611	1.7642230	2.3432756	.60333443 ,	7.9753527
V72	-.72591651	.46909461	-1.4475913	-1.7729517 ,	.32201740
V25	-1.1376109	.41985251	-2.7100395	-2.9376762 ,	-.23754357
V70	1.2626724	1.1948256	1.0575650	-1.2979293 ,	3.0230744
V14	4.2160475	3.3378606	1.2630964	-2.9429643 ,	11.375059
V23	-.95154077	.22211784	-3.8065044	-1.3374761 ,	-.38404541
V20	.56188879	.42074343	1.3731364	-.24061531 ,	1.1642917
V55	-.97126825	.50245435	-1.9322673	-2.0593602 ,	.10582966
V25	1.1775008	.36549620	3.1954779	.3571579 ,	1.9573336
V42	.57502243	.50387623	1.1401639	-.51712554 ,	1.6691735
V46	.30765237	1.2159410	.25320940	-2.2938571 ,	2.9155619
V41	-.98440641	.57130936	-1.7230737	-2.2087409 ,	.24092812
V7	-4.5944856	5.1279942	-.89603124	-15.5[illegible] ,	5.4031106
V92	.18591972	.36119548	.51727254	-.46906034 ,	.83191374
V91	-1.0074813	.37762323	-2.6721995	-1.5161155 ,	-.19584504
V8	.52275335	.21577420	2.4220684	.59954053E-01,	.98554363
V77	1.1739777	.54851896	2.1390932	-.31216933E-02,	2.3510772
V2	1.1716442	.77251386	1.5166561	-.4552443 ,	2.8255327
V18	-1.4197613	1.7205796	-.82512870	-5.1132597 ,	2.2737071
V79	-.46358303	.27965363	-1.6551649	-1.0613796 ,	.13361341
V12	-.68290578	.41842353	-1.6229125	-1.5446525 ,	.27886299
V74	-.67521557	.31070475	-2.1728255	-1.3446110 ,	-.11620102E-01
V33	.31204639	.71174929	.43841963	-1.2144393 ,	1.8395717
V75	1.7961437	.92469799	1.9424176	-.15713022 ,	3.7294295
V65	-.52879837	.37210757	-1.5570432	-1.3510591 ,	.27724321
V18	5.2054552	3.5771502	1.4552373	-2.4567283 ,	12.867719
V15	-4.1762549	3.0390226	-1.3731341	-10.797401 ,	2.4453910
V4	-.37708147	.31315724	-1.2094693	-1.0154600 ,	.34522792
V22	-.83241536	.91463256	-.91023782	-2.7379164 ,	1.1290353
V73	.34592343	.69321771	.49904120	-1.1515395 ,	1.8434433
V29	.20169020	.56478369	.35731391	-.77597415 ,	1.3014866
V13	.12589762	.32107770	.39720955	-.56024074 ,	.81233596
V42	-.17115696	.46972939	-.36344854	-1.2319623 ,	.89736063

SPSSSTEPWISE WITH 60 LEFT FROM CHOW   11/24/79   12.35.34.   PAGE 144

V3.	-.47245574	1.4152564	-.33759243	-3.5130449	, 2.5551334
V9.	.11155479	.01216017	.19417253	-1.1340932	, 1.4319177
CONSTANT	-5.7362908	2.1337275	-3.0763147	-11.432793	, -2.0397926

SPSSSTEPWISE WITH 56 LEFT FROM CHOW   11/24/79   12.35.34.   PAGE 148

FILE   NONAME   (CREATION DATE = 11/24/79)

* M U L T I P L E   R E G R E S S I O N *

DEPENDENT VARIABLE..   V98

SUMMARY TABLE

STEP	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	SIGNIFICANCE	MULTIPLE R	R SQUARE	R SQUARE CHANGE	SIMPLE R	OVERALL F	SIGNIFICANCE
1	V3		81.80622	0	.75704	.57312	.57312	.75704	81.83622	0
2	V63		15.04625	.000	.81151	.65870	.08559	.37304	57.93019	.000
3	V11		13.21893	.001	.84922	.72117	.05247	-.34276	50.85727	0
4	V64		9.53452	.003	.87209	.76054	.03935	.27552	46.05275	.000
5	V1		6.35425	.014	.88578	.78460	.02406	.06341	41.52583	0
6	V83		5.74671	.024	.89540	.80354	.01893	.49457	33.17360	.000
7	V53		4.37605	.041	.90414	.81662	.01408	.26202	35.31795	0
8	V95		3.44530	.067	.91062	.82905	.01193	-.01576	32.73534	.000
9	V27		3.09940	.084	.91589	.83850	.00914	.11209	30.57372	0
10	V15		3.73722	.059	.92153	.84931	.01082	.24105	25.30559	.000
11	V15		4.71542	.035	.92648	.86207	.01275	-.24230	23.37059	0
12	V17		4.15234	.047	.93417	.87267	.01060	.03153	25.55521	.000
13	V34		1.56429	.217	.93627	.87661	.00394	.10376	26.77775	0
14	V11		1.34132	.246	.93812	.88006	.00345	-.13757	25.15734	.000
15	V1		1.93267	.165	.94071	.88494	.00488	.49054	24.09319	0
16	V5		1.25435	.268	.94223	.88799	.00306	.22753	22.75303	.000
17	V99		2.64994	.096	.94551	.89475	.00576	.42313	22.50355	0
18	V23		2.33355	.134	.94871	.90005	.00530	.10463	22.01254	.000
19		V15	.00138	.971	.94871	.90005	-.00000	.24105	23.63472	0
20	V75		2.32605	.134	.95139	.90507	.00502	-.14503	23.30555	.000
21	V25		1.65427	.204	.95321	.90861	.00354	.07310	22.49530	0
22	V74		2.21249	.144	.95551	.91319	.00457	.51659	22.01224	.000
23	V14		2.43419	.123	.95820	.91814	.00496	.23524	21.49110	0
24	V23		3.60552	.065	.96172	.92491	.00577	-.09459	22.35458	.000
25	V20		22.15255	.000	.97547	.95212	.02721	-.03419	33.71511	0
26	V9		1.81319	.197	.97559	.95429	.00217	.18536	33.55614	.000
27	V4		2.12640	.153	.97815	.95673	.00248	.32471	32.75159	0
28	V55		1.63477	.181	.97924	.95891	.00213	.24647	32.35352	.000
29	V25		1.79113	.178	.98031	.96101	.00211	-.06292	31.95306	0
30	V42		1.65679	.205	.98124	.96283	.00182	-.05174	31.45241	.000
31	V49		1.35995	.252	.98199	.96431	.00147	-.03077	30.74159	0
32	V41		5.11015	.031	.98449	.96921	.00491	-.07568	33.37467	.000
33	V7		2.03217	.164	.98545	.97111	.00189	-.12923	33.61047	0
34	V42		1.25209	.272	.98603	.97226	.00116	.02226	32.95465	.000
35	V91		.65307	.427	.98634	.97287	.00061	.51879	31.51511	0
36	V6		1.13757	.295	.98688	.97393	.00105	.46526	30.76753	.000
37	V77		1.15343	.290	.98743	.97501	.00108	-.04439	30.09545	0
38	V2		.73346	.381	.98780	.97575	.00074	.71932	29.35562	.000
39	V13		.51058	.481	.98805	.97623	.00049	-.08709	27.75496	0
40	V70		.79090	.363	.98843	.97698	.00075	-.16089	26.81035	.000
41		V4	.00042	.984	.98843	.97698	-.00000	.52329	28.63136	0
42	V12		.67191	.421	.98876	.97762	.00063	.15377	27.58556	.000
43	V74		.63607	.433	.98905	.97822	.00060	-.13725	26.45750	0
44		V34	.00362	.960	.98905	.97822	-.00000	.10634	28.26260	.000
45	V33		1.61572	.207	.98980	.97971	.00149	.22509	28.47091	0
46	V75		.52101	.474	.99004	.98018	.00049	-.03114	27.20592	.000

SPSSSTEPWISE WITH 60 LEFT FROM CHOW 11/24/78 12.35.34. PAGE 149

FILE NONAME (CREATION DATE = 11/24/78)

47	V55		.45720	.502	.99526	.98062	.00043	.50396	25.91009	0
48	V14		.41258	.526	.99046	.98101	.00039	.56182	24.59559	.000
49	V15		.67728	.421	.99079	.95166	.00055	.24065	23.65161	0
50		V10	.59044	.983	.99079	.98166	-.00000	-.13327	25.46560	.000
51	V4		.97171	.337	.99124	.98255	.00095	.62579	24.95323	0
52		V16	.09255	.960	.99124	.98255	-.00000	-.24230	26.81285	.000
53	V22		.53637	.452	.99150	.98307	.00052	.60979	25.65473	0
54	V73		.59043	.452	.99177	.98361	.00054	-.08370	24.55420	.000
55		V9	.00157	.969	.99177	.98361	-.00590	.14538	26.51872	0
56	V23		.24858	.624	.99143	.98363	.00022	.07366	24.89670	.000
57	V13		.12473	.728	.99194	.98395	.00012	.13204	23.16251	0
58	V32		.06505	.817	.99157	.98401	.00005	.26073	21.40657	.000
59	V34		.07997	.781	.99201	.98405	.00008	.26715	19.74265	0
60	V34		.03770	.849	.99204	.98413	.00004	.10034	18.09196	.000

FILE NONAME (CREATION DATE = 11/24/78)

* M U L T I P L E R E G R E S S I O N *

OBSERVATION	Y VALUE	Y ESTIMATE	RESIDUAL	-250 0.0 +250
1.	-.5218189	-.7923779	.2705611	
2.	-.3657237	-.1167535	-.2489537	
3.	-.1737953	-.1532751	-.2052020E-01	
4.	1.808392	1.855541	-.5264603E-01	
5.	.1092316	.2748341E-01	.8174824E-01	
6.	-.3090736	-.2937563	-.1631036E-01	
7.	-2.960312	-2.943321	-.1659118E-01	
8.	-3.237127	-3.196753	-.4042339E-01	
9.	-.5589278	-.7760427	.1171142	
10.	-.7492139	-.7179732	-.3131555E-01	
11.	-2.546904	-2.557359	.1098441E-01	
12.	.5970722	.6015391	-.4516339E-02	
13.	-1.791625	-1.937533	.1520137	
14.	1.253361	1.337383	-.8432174E-01	
15.	1.101332	1.262287	-.1566949	
16.	-1.479100	-1.584215	.1051150	
17.	-2.350720	-2.298412	-.5230327E-01	
18.	-1.010135	-1.028112	.1017821E-01	
19.	2.165301	2.037363	.1470121	
20.	.2678142	.2792495	-.1142435E-01	
21.	-1.891302	-1.675521	-.2167507	
22.	.2872721	.1732117	.1050713	
23.	1.271359	1.250175	.6328037E-01	
24.	1.135419	1.029352	.1059157	
25.	-.7161315	-.7073185	-.9222063E-02	
26.	3.207364	3.253132	.2311205E-01	
27.	-.1052793	.2111762	-.3164539	
28.	1.138377	1.171557	-.3302413E-01	
29.	1.361328	1.767390	.2062222E-01	
30.	-.2022703	-.3227263	.1204550	
31.	-1.696116	-1.769415	.7329935E-01	
32.	.9203758	1.212785	-.2324052	
33.	-.3757430	-1.039520	.6377693E-01	
34.	-2.569179	-2.696307	.1274222	
35.	-2.356727	-2.414790	.5402382E-01	
36.	.2013082	.1274020	.7384013E-01	
37.	-.2684482	-.2971399	.2870234E-01	
38.	-1.556943	-1.595515	.3947215E-01	
39.	-2.280513	-2.159528	-.1209044	
40.	.8153307	.7425240	.3286324E-01	
41.	-.3255526	-.2999743	-.2556774E-01	
42.	1.341473	1.314245	.2722330E-01	
43.	.1239371	.1333495	-.9391432E-02	
44.	-.3841779	-.8114224	-.7275536E-01	
45.	.2663593	.2743203	-.1105301E-01	
46.	-5.066466	-4.592373	-.4676424	
47.	-.7257336	-.5756707	.1497711	
48.	-4.274165	-4.692373	.3261543	
49.	-2.757483	-2.700354	-.5692443E-01	
50.	-1.551365	-1.836367	.3640392E-01	
51.	.6611121	.5073413	.5772793E-01	
52.	.1339105	.4132363	-.3639742	

SPSSSTEPWISE WITH 62 LEFT FROM CHOW 11/24/78 12.35.34. PAGE 151

FILE NONAME (CREATION DATE = 11/24/78)

* M U L T I P L E   R E G R E S S I O N *

OBSERVATION	Y VALUE	Y ESTIMATE	RESIDUAL	-2SD	0.0	+2SD
53.	.5846350	.4763730	.1082770		I .	
54.	.6654479	.4574764	.1979311		I .	
55.	1.239573	1.067355	.1722147		I .	
56.	1.017690	1.057353	-.4063E032E-01		. I	
57.	-.2383154	-.1252433	-.1130732		. I	
58.	.1655574	.2644492	-.1520975E-01		.I	
59.	-.1924137	-.2511114	.5064758E-01		I .	
60.	1.515102	.9596307	.9654512		I	?
61.	.2592580	.8124090	-.5445509	.	I	
62.	.4712764	.9319301	-.4601215	.	I	
63.	-1.030672	-1.033795	.2512302E-01		I.	

NOTE - (*) INDICATES ESTIMATE CALCULATED WITH MEANS SUBSTITUTED
? INDICATES POINT OUT OF RANGE OF PLOT

NUMBER OF CASES PLOTTED 63.
NUMBER OF 2 S.D. OUTLIERS 1. OR 1.59 PERCENT OF THE TOTAL

VON NEUMANN RATIO 2.23711 DURBIN-WATSON TEST 2.25083

NUMBER OF POSITIVE RESIDUALS 33.
NUMBER OF NEGATIVE RESIDUALS 30.
NUMBER OF RUNS OF SIGNS 35.

EXPECTED NUMBER OF RUNS OF SIGNS 32.
EXPECTED S.D. OF RUN DISTRIBUTION 3.93742
UNIT NORMAL DEVIATE-
Z=(EXPECTED-OBSERVED)/S.D. .79205
PROBABILITY OF OBTAINING .GE. ABS(Z) .21709

99

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
BMD02R STEPWISE REGRESSION-REVISED
 HEALTH SCIENCES COMPUTING FACILITY, UCLA

 PROBLEM CODE                    SINGLE
 NUMBER OF CASES                     53
 NUMBER OF ORIGINAL VARIABLES        51
 NUMBER OF VARIABLES ADDED            0
 TOTAL NUMBER OF VARIABLES           51
 NUMBER OF SUB-PROBLEMS               1
 THE VARIABLE FORMAT IS    (7(12X,2F20.14/),12X,2F20.14/12X,3F20.14/12X,2F20.14/12X,3F20.14/12X,F20
                           .14/12X,F20.14/12X,3F20.14/6(20X/),32X,2F20.14/2(12X,3F20.14/),12X,F20.1
                           4/32X,2F20.14/12X,2F20.14/12X,F20.14/80X/32X,2F20.14/2(12X,3F20.14/),12X
                           ,F20.14,20X,F20.14/32X,F20.14)
```

VARIABLE	MEAN	STANDARD DEVIATION
1	4.77975	1.54137
2	3.70957	1.11632
3	3.10612	1.07592
4	3.10069	1.57344
5	3.78458	1.54491
6	5.24304	1.33939
7	.19345	.39543
8	.17373	.38375
9	.19048	.39543
10	.19048	.39583
11	.07349	.26561
12	.05349	.22583
13	.10349	.30583
14	.12069	.33263
15	.22952	.39553
16	.31270	.46527
17	.10349	.34530
18	.26508	.44532
19	.31745	.46323
20	1.74483	3.51585
21	1.19943	2.31271
22	.72407	1.14223
23	1.02978	2.22154
24	.72603	1.98675
25	1.19371	2.92492
26	1.03523	2.41395
27	.42215	1.13723
28	.70540	1.70513
29	.36635	1.51171
30	1.59765	3.53281
31	1.77555	2.86913
32	.50253	1.42441
33	1.31316	2.76955
34	.52766	1.35517
35	1.13613	2.47545
36	.24973	1.50439
37	1.72400	3.33751
38	.38085	2.30781
39	.22916	1.15722
40	.35709	2.50277
41	.38326	1.54183

Figure 8. Selected Output from BMD

42	.77127	2.07571
43	2.[illegible]	3.5[illegible]
44	3.[illegible]	2.8[illegible]
45	2.70926	1.58[illegible]
46	3.72514	4.[illegible]
47	2.72012	3.32032
48	1.27600	1.58707
49	2.77773	2.37411
50	2.18994	2.72105
51	3.20155	4.41715
52	2.75495	3.62822
53	1.75278	1.27785
54	2.14316	3.25031
55	1.76556	1.93342
56	2.31349	2.74525
57	2.37532	4.57792
58	2.56593	3.12050
59	.34975	1.53394
60	1.78930	2.05934
61	-.77634	1.54952

SUB-PROBLEM 1

DEPENDENT VARIABLE	61
MAXIMUM NUMBER OF STEPS	122
F-LEVEL FOR INCLUSION	.010000
F-LEVEL FOR DELETION	.005000
TOLERANCE LEVEL	.001000

STEP NUMBER 1
VARIABLE ENTERED 3

MULTIPLE R	.7570
STD. ERROR OF EST.	1.0464

ANALYSIS OF VARIANCE

	DF	SUM OF SQUARES	MEAN SQUARE	F RATIO
REGRESSION	1	89.665	89.665	81.896
RESIDUAL	61	66.787	1.095	

VARIABLES IN EQUATION

VARIABLE	COEFFICIENT	STD. ERROR	F TO REMOVE
(CONSTANT	-3.37339721	)	
3	1.11166	.12361	81.8962 (2)

VARIABLES NOT IN EQUATION

VARIABLE	PARTIAL CORR.	TOLERANCE	F TO ENTER
1	.10129	.6535	.0224 (2)
2	.02626	.1112	.0414 (2)
4	.39947	.7663	11.3245 (2)
5	.27227	.9965	4.8956 (2)
6	.02572	.1431	.0397 (2)
7	-.00629	.9763	.0030 (2)
8	.23694	.5539	3.5771 (2)
9	.12664	.9214	.9711 (2)
10	-.29504	.9941	5.7254 (2)
11	-.44476	.3811	14.7352 (2)
12	.21316	.6983	2.7935 (2)
13	.03528	.9700	.1324 (2)
14	.25696	.5911	4.2100 (2)
15	.42722	.9974	13.3959 (2)
16	-.24165	.5973	3.7256 (2)
17	-.01215	.5993	.3015 (2)
18	.25217	.7751	4.0552 (2)
19	-.03759	.5126	.0536 (2)
20	.02412	.7325	.0357 (2)
21	.00954	.7353	.0359 (2)
22	.00359	.4994	.0007 (2)
23	-.02941	.9961	.0534 (2)
24	.03510	.9371	.0770 (2)
25	.20465	.5537	2.5173 (2)
26	.19573	.9953	2.1577 (2)
27	.17225	.9792	1.8575 (2)
28	.21521	.9613	2.5337 (2)
29	.19519	.9970	2.3746 (2)
30	.13751	.9744	2.3315 (2)
31	.13561	.9566	1.1411 (2)
32	.15753	.9516	1.5372 (2)
33	-.30937	.9747	6.3590 (2)

.	34	-.31461	.5500	6.5035 (2)
.	35	-.24114	[illegible]	3.7044 (2)
.	36	-.13223	[illegible]	2.5683 (2)
.	37	.25341	[illegible]	4.5581 (2)
.	38	.25360	[illegible]	4.4736 (2)
.	39	.27355	.5924	[illegible]
.	40	.25910	[illegible]	4.3180 (2)
.	41	.22545	.5513	4.2241 (2)
.	42	.25670	.9863	4.3636 (2)
.	43	.44776	[illegible]	15.0462 (2)
.	44	.40301	.7213	11.5432 (2)
.	45	.39125	.7229	10.5447 (2)
.	46	-.15542	.3956	1.7032 (2)
.	47	-.20550	[illegible]	2.5157 (2)
.	48	-.17211	.8865	1.5310 (2)
.	49	-.04931	.5567	[illegible] (2)
.	50	-.15531	.6373	2.1470 (2)
.	51	.24865	[illegible]	3.5806 (2)
.	52	.23974	.7411	3.3455 (2)
.	53	.22216	.7155	3.1132 (2)
.	54	.25325	.5525	4.2860 (2)
.	55	.21017	.4455	2.7725 (2)
.	56	.23153	[illegible]	3.3365 (2)
.	57	[illegible]	[illegible]	.0060 (2)
.	58	-.03474	.5964	.0725 (2)
.	59	-.03719	.5731	.0827 (2)
.	60	-.06535	.5953	.2611 (2)

STEP NUMBER 2
VARIABLE ENTERED 43

MULTIPLE R .8116
STD. ERROR OF EST. .9434

ANALYSIS OF VARIANCE

	DF	SUM OF SQUARES	MEAN SQUARE	F RATIO
REGRESSION	2	103.155	51.575	57.900
RESIDUAL	60	53.396	.890	

VARIABLES IN EQUATION					VARIABLES NOT IN EQUATION			
VARIABLE	COEFFICIENT	STD. ERROR	F TO REMOVE	.	VARIABLE	PARTIAL CORR.	TOLERANCE	F TO ENTER
(CONSTANT	-4.62472477	)		.				
3	1.07143	.11211	91.3258 (2)	.	1	.11656	.6336	.8415 (2)
43	.12758	.03289	15.0462 (2)	.	2	.03214	.1118	.0603 (2)
				.	4	.25546	.3021	4.3202 (2)
				.	5	.11527	[illegible]	.7980 (2)
				.	6	.12916	.4153	1.0065 (2)
				.	7	-.16149	.7751	1.5726 (2)
				.	8	.27565	.3435	4.8736 (2)
				.	9	.06527	[illegible]	.2524 (2)
				.	10	-.05826	.3803	.2044 (2)
				.	11	-.42763	.9732	13.2169 (2)
				.	12	.15420	.5771	1.5348 (2)

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	MULTIPLE R	RSQ	INCREASE IN RSQ	F VALUE TO ENTER OR REMOVE	NUMBER OF INDEPENDENT VARIABLES INCLUDED
1	3		.7579	.5731	.5731	81.[illegible]	1
2	43		.8116	.6337	.0655	17.0462	2
3	11		.8432	.7212	.0625	17.2159	3
4	39		.8721	.7505	.0294	[illegible]	4
5	8		.8858	.7445	.0241	[illegible]	5
6	54		.8904	.8235	.0159	[illegible]	6
7	36		.9024	.8130	.0145	[illegible]	7
8	60		.9125	.8211	.0110	3.[illegible]	8
9	26		.9177	.8335	.0094	2.1364	9
10	15		.9216	.8432	.0106	3.[illegible]322	10
11	16		.9235	.8521	.0128	[illegible]	11
12	17		.9312	.8727	.0105	4.1524	12
13	59		.9363	.8755	.0039	1.[illegible]	13
14	10		.9371	.8801	.0035	1.7314	14
15	1		.9407	.8843	.0043	1.8321	15
16	5		.9423	.8830	.0031	1.2[illegible]	16
17	55		.9459	.8943	.0055	2.[illegible]	17
18	27		.9487	.9021	.0053	2.3336	18
19		15	.9477	.9000	-.0001	.1014	17
20	47		.9514	.9051	.0050	2.3270	18
21	25		.9572	.9105	.0035	1.[illegible]	19
22	45		.9576	.9132	.0045	2.212[illegible]	20
23	14		.9562	.9131	.0050	2.[illegible]2	21
24	23		.9617	.9243	.0055	[illegible]	22
25	20		.9758	.9521	.0272	22.1525	23
26	9		.9763	.9543	.0022	1.3052	24
27	4		.9772	.9559	.0026	2.1254	25
28	40		.9792	.9519	.0021	1.[illegible]	26
29	24		.9803	.9610	.0021	1.3911	27
30	36		.9812	.9629	.0019	1.[illegible]	28
31	34		.9820	.9643	.0015	1.35[illegible]	29
32	35		.9845	.9692	.0001	[illegible].1302	30
33	7		.9814	.9711	.0019	2.[illegible]	31
34	57		.9860	.9723	.0012	1.2[illegible]21	32
35	29		.9867	.9735	.0012	1.35[illegible]	33
36	6		.9873	.9746	.0013	1.[illegible]	34
37	56		.9877	.9755	.0005	.3121	35
38	49		.9884	.9770	.0014	1.[illegible]	36
39	28		.9890	.9782	.0012	1.[illegible]	37
40	2		.9894	.9785	.0007	.7[illegible]	38
41		4	.9894	.9783	-.0000	.0007	37
42	57		.9895	.9792	.0003	[illegible]	38
43	12		.9898	.9795	.0005	.6519	39
44	31		.9901	.9803	.0003	.5753	40
45	46		.9904	.9808	.0005	[illegible]	41
46	19		.9909	.9818	.0011	1.[illegible]2	42
47	48		.9911	.9822	.0003	.3513	43
48	50		.9915	.9830	.0005	[illegible]	44
49	15		.9917	.9834	.0004	[illegible]	45
50		52	.9917	.9834	-.0000	[illegible]	44
51	30		.9918	.9837	.0003	[illegible]	45
52		25	.9918	.9837	-.0000	[illegible]	44
53		3	.9919	.9837	-.0000	[illegible]	43
54	51		.9919	.9839	.0002	[illegible]	44

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
PROGRAM MULTI       74/74    OPT=1                           FTN 4.6+446          11/30/78  18.35.40         PAGE     1

                 PROGRAM MULTI(INPUT,OUTPUT,TAPE1,TAPE3)
          C      GROUP1 WITH GROUP2AND 3 SELECTED BY ROAR
          C
          C      VARIABLE NO.   VARIABLE NAME
          C          1              UP
          C          2              W
          C          3              DO
          C          4              %SS
          C          5              NR
          C          6              SF
          C          7              ER
          C          8              DIG
          C          9              FM
          C         10              MATR
          C         11              NF*UP
          C         12              NF*W
          C         13              NF*DO
          C         14              NF*PD
          C         15              ND*UP
          C         16              ND*V
          C         17              EF*DO
          C         18              DIG*UP
          C         19              DIG*V
          C         20              DIG*W
          C         21              DIG*%SS
          C         22              DIG*PD
          C         23              INT
          C         24              FM*W
          C         25              CM*PD
          C         26              PF*W
          C         27              PF*%SS
          C
                 DIMENSION RR(19361),IJOB(4),IXS(31),STAT(50),IXV(31)
                 DIMENSION IXB(31),BEST(465,4),NVAR(300)
                 DIMENSION IW(3174),WK(310)
                 DIMENSION YM(31),NBR(6),XX(100,31),TEMP(31),VC(31,31)
          C
          C           TAPE1 IS DATA TAPE.
          C      DATA IS IN 4F20.10 FORMAT
          C      INDEPENDENT VARIABLES ARE FIRST
          C      DEPENDENT VARIABLE IS LAST
          C           TAPE3 IS CONTROL TAPE
          C      NUMBER OF INDEPENDENT VARIABLES
          C      NUMBER OF OBSERVATIONS
          C      IN I3 FORMAT
          C
                 IX=100
          C
          C      THIS IS THE MAXIMUM NUMBER OF DATA POINTS, AND
          C      SHOULD BE CHANGED IF DIMENSIONS ARE CHANGED.
          C
                 PRINT 2000
          2000   FORMAT(" MAXIMUM NUMBER OF VARIABLES = 29")
                 READ (3,3000)IVAR
          3000   FORMAT(I3)
                 PRINT 2001,IVAR
          2001   FORMAT(" THERE ARE ",I3," INDEPENDENT VARIABLES")
```

Figure 9. Sample Leaps and Bounds

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
      NBR(1)=IVAR+1
      NBR1=NBR(1)
      PRINT 2002
 2002 FORMAT(" MAXIMUM NUMBER OF OBSERVATIONS = 100")
      READ (7,3000)NBR(2)
      NBR2=NBR(2)
      PRINT 2003,NBR(2)
 2003 FORMAT(" NUMBER OF OBSERVATIONS = ",I3)
      PRINT 2004
 2004 FORMAT(10X,"INPUT MATRIX")
      DO 1 I=1,NBR2
      READ(1,1000)(X(I,J),J=1,NBR1)
 1    CONTINUE
 1000 FORMAT(12X,F20.14,21X,F20.14/12X,2F20.14/32X,F20.14/12X,2F20.14/
     %32X,F20.14/12X,F20.14,20X,F20.14/32X,F20.14/12X,2F20.14/12X,F20.14
     $/20X/32X,2F20.14/2(80X/),32X,F20.14/5(80X/),32X,2F20.14/12X,
     *F20.14,20X,F20.14/12X,F20.14/12X,F20.14/30X/12X,F20.14/12X,F20.14/
     $4(80X/),12X,F20.14,20X,F20.14/32X,F20.14)
      IF(NBR(2).GT.IX)STOP1
      NBR(3)=NBR(2)
      NBR(4)=NBR(5)=NBR(6)=1
C
      CALL BECOVM(XX,IX,NBR,TEMP,XM,RR,IER)
C
      PRINT 2006
 2006 FORMAT(/10X,"SUM OF SQUARES AND CROSS PRODUCTS MATRIX"/)
C
      CALL VCVTSF(RR,NBR(1),VC,27)
C
      PRINT 2007,((VC(I,J),J=1,NBR1),I=1,NBR1)
 2007 FORMAT(5G20.10)
      PRINT 2008
 2008 FORMAT(/" MEAN"/)
      PRINT 2007,(XM(I),I=1,NBR1)
      IF(IER.NE.0)PRINT 2009,IER
 2009 FORMAT(/" ***********************"/" ERROR IN BECOVM"/" ERROR NUMBER
     1",I3/" ***********************"/)
      K7=NBR(1)
      PRINT 2010,K7
 2010 FORMAT(/" NUMBER OF VARIABLES = ",I3)
      IJOB(1)=NBR(2)-1
      IF(IJOB(1).LT.K7)STOP2
      IJOB(2)=1
 3    CONTINUE
      IJOB(2)=IJOB(2)+1
      IF(IJOB(2).GT.3)STOP3
      IJOB(3)=1
      IJOB(4)=2
      IR=300
      KX=K7-1
C
      CALL RLEAP(RR,K7,IJOB,IXS,STAT,IXV,NVAR,IX3,BEST,IR,WK,IW,IER)
C
      IF(IER.NE.0)PRINT 2011,IER
 2011 FORMAT(/" ***********************"/" ERROR IN RLEAP"/" ERROR NUMBER "
     1,I3/" ***********************"/)
      IF(IER.GT.128) GO TO 3
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
            IF(IER.NE.37) GO TO 5
            PRINT 2012,(IW(I),I=1,KX)
      2012  FORMAT(" VARIABLES DELETED = ",20I5)
            ICNT=0
            DO 4 I=1,KX
            IF(IW(I).EQ.0) ICNT=ICNT+1
      4     CONTINUE
            K7=ICNT+1
      C
      5     CALL USLEAP(IJOB,K7,IXS,STAT,IXV,NVAR,IX3,BEST,IB)
      C
            IF(IER.NE.0)PRINT 2013,IER
            GO TO 3
      2013  FORMAT(/" ********************"/" ERROR IN USLEAP"/" ERROR NUMBER
           1 ",I3/" ********************"/)
            END
```

```
REGRESSIONS WITH   1 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .556119E+02        2
      .355459E+02        3

REGRESSIONS WITH   2 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .645144E+02        2  7
      .628925E+02        2  3

REGRESSIONS WITH   3 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .695642E+02        2  6  7
      .634315E+02        2  6 23

REGRESSIONS WITH   4 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .749402E+02        2  6  7 20
      .745625E+02        2  6  7 19

REGRESSIONS WITH   5 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .755431E+02        2  6  7 18 20
      .761309E+02        2  6  7  9 20

REGRESSIONS WITH   6 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .724033E+02        2  6  7  9 19 20
      .730849E+02        2  6  7 19 20 21

REGRESSIONS WITH   7 VARIABLE(S) (ADJUSTED R-SQUARED)

     CRITERION          VARIABLES
      .734551E+02        1  2  6  7  9 19 20
      .799199E+02        2  6  7  9 18 20 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
REGRESSIONS WITH   8 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .837733E+02         1  2  6  7  9 19 21 27
        .832952E+02         1  2  5  6  7  9 19 20

REGRESSIONS WITH   9 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .821573E+02         1  2  6  7  9 19 21 26 27
        .814034E+02         1  2  5  6  7  9 19 20 27

REGRESSIONS WITH  10 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .833316E+02         1  2  5  7  9 19 21 23 26 27
        .835948E+02         1  2  5  6  7  9 19 20 26 27

REGRESSIONS WITH  11 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .835908E+02         1  2  5  6  7  9 19 20 23 26 27
        .835675E+02         1  2  6  7  9 19 20 22 23 26 27

REGRESSIONS WITH  12 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .849494E+02         1  2  6  7 13 14 19 20 24 25 26 27
        .844933E+02         1  2  5  6  7  9 19 20 22 23 26 27

REGRESSIONS WITH  13 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .853630E+02         1  2  6  7 13 14 19 20 23 24 25 26 27
        .850549E+02         1  2  7 13 14 18 20 22 23 24 25 26 27

REGRESSIONS WITH  14 VARIABLE(S) (ADJUSTED R-SQUARED)

       CRITERION           VARIABLES
        .854055E+02         1  2  6  7 13 14 19 20 22 23 24 25 26 27
        .871803E+02         1  2  6  7 13 14 18 19 20 23 24 25 26 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
REGRESSIONS WITH  15 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8937975E+02       1  2  6  7 13 14 18 19 20 22 23 24 25 26 27
    .8919745E+02       1  2  6  7 13 14 15 18 20 22 23 24 25 26 27

REGRESSIONS WITH  16 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8957415E+02       1  2  6  7 13 14 17 18 19 20 22 23 24 25 26 27
    .8363333E+02       1  2  6  7 12 13 14 15 16 18 22 23 24 25 26 27

REGRESSIONS WITH  17 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8933576E+02       1  2  6  7 12 13 14 15 16 18 20 22 23 24 25 26 27
    .8831146E+02       1  2  6  7 12 13 14 15 16 17 19 22 23 24 25 26 27

REGRESSIONS WITH  18 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8934085E+02       1  2  6  7 12 13 14 15 16 17 18 20 22 23 24 25 26 27
    .8931446E+02       1  2  6  7 12 13 14 15 16 17 18 19 22 23 24 25 26 27

REGRESSIONS WITH  19 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8917275E+02       1  2  5  6  7 12 13 14 15 16 17 18 20 22 23 24 25 26 27
    .8910346E+02       1  2  6  7 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27

REGRESSIONS WITH  20 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8901345E+02       1  2  6  7  8 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
    .8900255E+02       1  2  6  7 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  21 VARIABLE(S) (ADJUSTED R-SQUARED)

    CRITERION          VARIABLES
    .8937975E+02       1  2  5  6  7  8 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
    .8936336E+02       1  2  5  6  7 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
```

```
REGRESSIONS WITH  22 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8958375E+02       1  2  5  6  7  9 11 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
         .8957346E+02       1  2  5  6  7 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  23 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8949715E+02       1  2  4  6  7  8  9 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
         .8947546E+02       1  2  3  5  6  7  9 10 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27

REGRESSIONS WITH  24 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8933736E+02       1  2  3  5  6  7  9 10 11 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
         .8932516E+02       1  2  3  5  6  7 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  25 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8913935E+02       1  2  3  5  6  7  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

         .8907735E+02       1  2  3  5  6  7  8 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27


REGRESSIONS WITH  26 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8890675E+02       1  2  3  4  5  6  7  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
                           27
         .8865125E+02       1  2  3  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
                           27

REGRESSIONS WITH  27 VARIABLE(S) (ADJUSTED R-SQUARED)

        CRITERION          VARIABLES
         .8659798E+02       1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
                           26 27
```

BEST REGRESSIONS WITH 20 VARIABLE(S) (ADJUSTED R-SQUARED)

VARIABLE	COEFFICIENT	PARTIAL F	ALPHA
1	.713871E+00	.144190E+02	.450352E-03
2	.359494E+00	.385712E+01	.122545E-01
5	-.287899E+01	.143170E+02	.483531E-03
7	-.219569E+01	.791070E+02	.171434E-06
9	.430154E+01	.459592E+01	.323934E-01
12	-.343076E+00	.210727E+01	.154225E+00
13	-.470354E+00	.159317E+02	.260250E-03
14	.677722E+00	.145130E+02	.433744E-03
15	.274324E+00	.303324E+01	.903310E-01
16	-.292464E+00	.733734E+01	.337390E-01
17	.293229E+00	.629130E+01	.160123E-01
18	-.390123E+00	.117268E+02	.138749E-02
19	-.971578E+00	.224170E+01	.141235E+00
20	.267532E+01	.434745E+01	.315395E-01
22	.551307E+00	.255784E+01	.119173E+00
23	.239272E+00	.397725E+01	.294133E-02
24	.700385E+00	.134343E+02	.587959E-03
25	-.540675E+00	.216154E+02	.466115E-04
26	.629916E+00	.270533E+02	.550300E-05
27	-.786378E+00	.331529E+02	.220500E-06

```
REGRESSIONS WITH   1 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .1731215E+03        2
         .2607555E+03        3

REGRESSIONS WITH   2 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .1232056E+03        2  7
         .1352655E+03        2  7

REGRESSIONS WITH   3 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .1024395E+03        2  6  7
         .1031755E+03        2  7 23

REGRESSIONS WITH   4 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .7447325E+02        2  6  7 20
         .7639435E+02        2  6  7 19

REGRESSIONS WITH   5 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .6377775E+02        2  6  7 18 20
         .5837355E+02        2  6  7  9 20

REGRESSIONS WITH   6 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .5711375E+02        2  6  7  9 15 25
         .5863745E+02        2  6  7 13 20 24

REGRESSIONS WITH   7 VARIABLE(S) (MALLOWS  CP)

        CRITERION          VARIABLES
         .5203715E+02        1  2  6  7  9 15 20
         .5420525E+02        2  6  7  9 15 20 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
REGRESSIONS WITH   8 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .459572E+02         1  2  6  7  9 15 21 27
        .437209E+02         1  2  5  6  7  9 13 20

REGRESSIONS WITH   9 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .393366E+02         1  2  6  7  9 15 21 26 27
        .434414E+02         1  2  5  6  7  9 13 20 27

REGRESSIONS WITH  10 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .353351E+02         1  2  6  7  9 15 20 23 26 27
        .371937E+02         2  6  7 13 14 15 21 25 26 27

REGRESSIONS WITH  11 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .339256E+02         1  2  6  7 13 14 15 20 25 26 27
        .333459E+02         1  2  5  6  7  9 13 20 23 26 27

REGRESSIONS WITH  12 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .219307E+02         1  6  7 13 14 15 22 23 24 25 26 27
        .243017E+02         2  3  5  6  7 11 13 19 20 24 26 27

REGRESSIONS WITH  13 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .173646E+02         1  2  6  7 13 14 19 22 23 24 25 25 27
        .207335E+02         1  6  7 13 14 17 13 22 23 24 25 25 27

REGRESSIONS WITH  14 VARIABLE(S) (MALLOWS  CP)

       CRITERION           VARIABLES
        .153050E+02         1  2  5  7 13 14 19 20 22 23 24 25 26 27
        .179710E+02         1  2  6  7 12 13 14 18 22 23 24 25 26 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
REGRESSIONS WITH  15 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .153146E+02        1  2  6  7 13 14 17 18 20 22 23 24 25 25 27
       .153258E+02        1  2  6  7 12 13 14 15 20 22 23 24 25 25 27

REGRESSIONS WITH  16 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .132511E+02        1  2  6  7  9 13 14 17 18 19 20 23 24 25 26 27
       .132843E+02        1  2  6  7 13 14 17 18 19 20 21 23 24 25 26 27

REGRESSIONS WITH  17 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .141932E+02        1  2  6  7  9 13 14 17 18 19 21 22 23 24 25 26 27
       .142371E+02        1  2  6  7 13 14 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  18 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .152452E+02        1  2  6  7  9 13 14 15 16 17 18 19 20 23 24 25 26 27
       .153074E+02        1  2  6  7 13 14 15 16 17 18 19 20 21 23 24 25 26 27

REGRESSIONS WITH  19 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .156226E+02        1  2  6  7  8 13 14 15 16 17 18 19 20 22 23 24 25 26 27
       .158713E+02        1  2  6  7 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  20 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .157358E+02        1  2  6  7  9 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
       .158250E+02        1  2  6  7 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  21 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .173955E+02        1  2  5  6  7  9 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
       .174251E+02        1  2  5  6  7 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

```
REGRESSIONS WITH  22 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .192859E+02        1  2  4  6  7  9 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
       .203721E+02        1  2  4  5  7  8  9 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  23 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .209241E+02        1  2  4  6  7  8  9 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
       .211445E+02        1  2  4  5  6  7  9 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  24 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .229335E+02        1  2  3  5  6  7  9 10 11 12 13 14 15 16 17 18 19 20 22 23 24 25 26 27
       .229420E+02        1  2  4  5  6  7  8  9 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  25 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .242423E+02        1  2  3  5  6  7  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

       .244553E+02        1  2  3  5  5  7  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

REGRESSIONS WITH  26 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .250134E+02        1  2  3  4  5  6  7  8 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
                         27
       .252001E+02        1  2  3  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
                         27

REGRESSIONS WITH  27 VARIABLE(S) (MALLOWS  CP)

      CRITERION          VARIABLES
       .233300E+02        1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
                         26 27
```

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

BEST REGRESSIONS WITH 16 VARIABLE(S) (MALLOWS CP)

VARIABLE	COEFFICIENT	PARTIAL F	ALPHA
1	.3144195E+00	.2031655E+02	.4434135E-04
2	.3326135E+00	.7315[illegible]F+01	.354[illegible]35E-02
6	-.32[illegible]3[illegible]6E+01	.13[illegible]571E+02	.[illegible]573[illegible]E-04
7	-.2[illegible]2513[illegible]+01	.[illegible]3[illegible]1[illegible]1E+02	.2917135E-08
8	.[illegible]193[illegible]0E+01	.[illegible]3[illegible]0E+01	.[illegible]65[illegible]33E-02
13	-.[illegible]024175E+00	.1575095E+02	.793212E-04
14	.47[illegible]2[illegible]9E+00	.12[illegible]1[illegible]95E+02	.33[illegible]3[illegible]E-03
17	.[illegible]E+00	.337[illegible]06E+01	.4395145E-02
13	-.[illegible]1555E+10	.[illegible]0[illegible]29E+01	.4931575E-02
19	-.167[illegible]E+01	.[illegible]72[illegible]67E+01	.317[illegible]E-02
20	.[illegible]E+01	.[illegible]51E+02	.15[illegible]115E-04
23	.1[illegible]4[illegible]1E+00	.[illegible]3[illegible]E+01	.[illegible]375[illegible]5E-02
24	.[illegible]173[illegible]E+00	.122[illegible]03[illegible]E+02	.104[illegible]5E-02
25	-.4170[illegible]3E+00	.71[illegible]103E+02	.[illegible]41[illegible]72E-04
25	.7315[illegible]6E+00	.2[illegible]56[illegible]E+02	.1052[illegible]E-04
27	-.53321[illegible]E+00	.[illegible]497[illegible]E+02	.3913[illegible]E-06

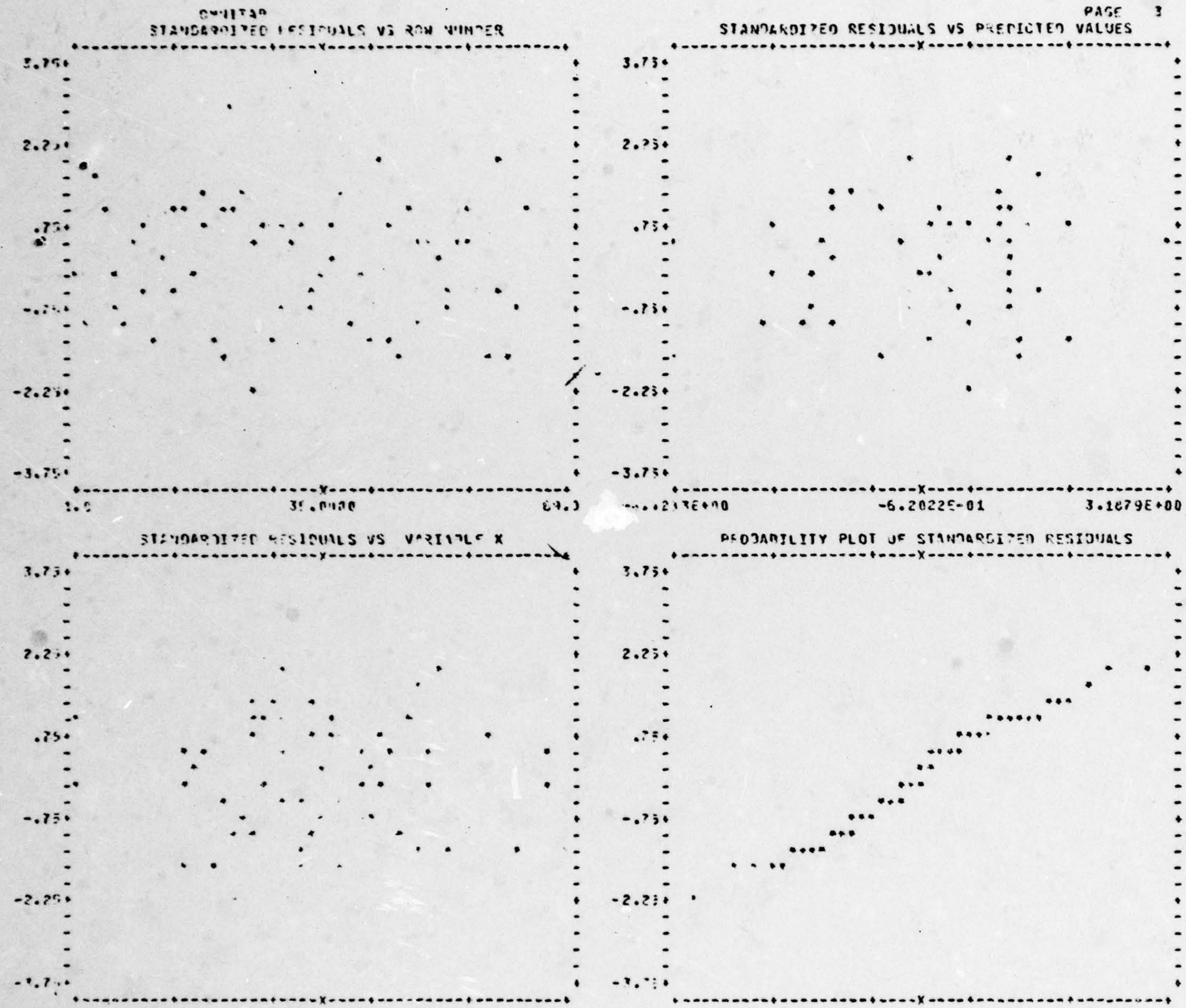
.... LJPOOFL //// END OF LIST ////

```
                    LEAST SQUARES FIT FOR      DATA      IN COLUMN   22
              AS A LINEAR FUNCTION OF 21 PREDICTOR VARIABLES IN COLUMNS   1,   2,   3,   4,   5,   6,   7,   8,
 9,  10,  11,  12,  13,  14,  15,  16,  17,  18,  19,  20,  21
                 USING  43 NON-ZERO WEIGHTS AND    5 ZERO WEIGHTS IN COLUMN   23
```

ROW	PREDICTOR VARIABLES IN COL. 2	COL. 3	COL. 4	DATA COL. 22	PREDICTED VALUES	STD. DEV. OF PRED. VALUES	RESIDUALS	STD. RES.	WEIGHTS
1	9.327	2.874	0.	-.32151600	-.32052772	.19050934	-3.3828332-0	-.00	1.000
2	10.52	3.660	0.	-.36172700	-.33361213	.30553105	-.03205297	-.06	1.000
3	9.307	3.609	0.	-.17779500	.27167103	.25650963	-.44942109	-1.03	1.000
4	11.22	3.401	0.	1.8079920	1.2769235	.39537633	.53207113	1.75	1.000
5	9.037	3.665	0.	.16123200	-.21333325	.29509078	.37241226	.92	1.000
6	7.815	3.457	0.	-.80057600	-.54259401	.27169439	-.25791059	-.61	1.000
7	5.732	1.366	0.	-2.9065120	-2.8935377	.29191265	-.01288309	-.04	1.000
8	7.125	.2231	0.	-3.2371775	-3.0012346	.47654279	-.23564235	-.91	1.000
9	12.95	2.634	0.	-.56192300	-.81273195	.29305795	.21340938	.55	1.000
10	9.215	2.572	0.	-.74528900	-.55303395	.24457593	-.08955101	-.72	1.000
11	[illegible]	1.972	0.	-2.5409000	-2.8280722	.30578913	.23117416	.71	1.000
12	13.92	3.745	0.	.59707200	1.0234733	.35125391	-.42035339	-1.22	1.000
13	7.053	2.915	0.	-1.7815256	-1.3764234	.23152463	.09474352	.21	1.000
14	7.317	4.094	0.	1.2170015	.35663529	.15551505	.44503271	.95	1.000
15	3.637	4.500	0.	1.1013620	1.2159372	.17437229	-.18757115	-.40	1.000
16	5.373	.3975	0.	-1.4751093	-1.3491919	.22371314	.47009130	1.05	1.000
17	5.463	.1623	0.	-2.3577200	-2.2439533	.27462824	-.10576411	-.25	1.000
18	7.221	2.711	0.	-1.0211350	-1.5337075	.27467177	.52317252	1.38	1.000
19	9.537	4.771	0.	2.1919019	1.3196230	.16179155	.37737209	.89	1.000
20	9.362	4.056	0.	.20741100	.82917430	.22358570	-.62135436	-1.47	1.000
21	7.265	1.224	0.	-1.3910920	-1.1082379	.26105645	-.71076403	-1.71	1.000
22	7.133	2.442	0.	.38127300	-.19435335	.12612155	.47183336	.97	1.000
23	3.031	4.566	0.	1.2713899	.73802433	.10562375	.53213451	1.15	1.000
24	9.215	4.472	0.	1.1374032	.54671037	.25533493	.48873733	1.21	1.000
25	7.625	4.149	0.	-.71463100	.24505471	.29723020	-.95103435	-2.37	1.000
26	12.44	4.317	0.	3.2620540	3.1514535	.15636407	.09423043	.51	1.000
27	9.743	2.519	0.	-.16127100	-.44564523	.23717343	.34113328	.77	1.000
28	10.43	3.446	0.	1.1748730	1.5776233	.30492397	-.53303033	-1.35	1.000
29	11.72	3.164	0.	1.3611260	1.1377415	.33405704	.23023735	.62	1.000
30	9.221	3.474	0.	-.20327100	.12917395	.32222493	-.23147037	-.63	1.000
31	7.147	2.345	0.	-1.4911150	-2.0064443	.23330018	.14277373	.43	1.000
32	7.416	4.304	0.	.09838300	.53009151	.20028433	.30031649	.74	1.000
33	7.324	2.379	0.	-.13774300	-.72591114	.25563503	-.24917136	-.53	1.000
34	7.034	2.137	0.	-2.4661750	-2.2512734	.31551350	-.27758364	-.72	1.000
35	3.917	1.346	1.000	-2.3517270	-2.2815455	.22105175	-.09119216	-.22	1.000
36	9.222	7.777	1.000	.20130300	-.35461512	.37577265	.24556451	.65	1.000
37	9.243	4.713	1.000	-.26846600	-.28732954	.27861563	.01534542	.04	1.000
38	3.743	2.358	1.000	-1.6660432	-1.3721335	.34250640	.41633055	1.33	1.000
39	7.643	2.706	1.000	-2.2805130	-1.5092641	.29113664	-.38130859	-.91	1.000
40	3.427	3.219	1.000	.31539300	.31753175	.41672449	-.00233117579	-.01	1.000
41	12.37	3.714	1.000	-.32454300	-.32450332	.40947646	-.00113234763	-.00	1.000
42	11.46	5.157	1.000	1.7411730	1.7192327	.31222694	-.43776759	-1.21	1.000
43	11.72	7.345	1.000	.18739500	-.55007755	.24913136	.57403356	1.39	1.000
44	11.53	4.705	1.000	-.39117500	-.30468483	.40535577	-.43951772	-1.43	1.000
45	9.153	3.567	0.	.20433500	.15751327	.47102222	.10445183	.70	1.000
46	5.727	2.145	0.	-4.3014650	-4.4282171	.39635115	-.56117734	-1.61	1.000
47	11.72	7.777	0.	-.72126500	-1.1646743	.30957275	.43322145	1.16	1.000
48	5.727	2.143	0.	-4.2741230	-4.4242371	.30625155	.15010325	.31	1.000

Figure 10. Sample OMNITAB Output

50	17.03	2.4[illegible]	1.000	-1.[illegible]19640	-2.350[illegible]39	.24332042	.208[illegible]491	.43	1.000
51	[illegible]	2.[illegible]	0.	.56211283	.30912723	.12500264	.0[illegible]1[illegible]4166	.11	1.000
52	3.[illegible]1	3.[illegible]	0.	.[illegible]3[illegible]1080	.30912783	.12[illegible]263	-.36975743	-.76	1.000
53	3.137	3.3[illegible]	0.	.5[illegible]63000	.70039381	.13052636	-.15626031	-.41	1.000
54	3.[illegible]	2.171	0.	.58[illegible]60000	.45723254	.15590612	.21817572	.47	1.000
55	3.250	3.592	0.	1.23[illegible]700	.756[illegible]331	.14109014	.4[illegible]9[illegible]502	.90	1.000
56	4.2[illegible]	3.5[illegible]	0.	1.3175000	.73405731	.14109014	.22737192	.46	1.000
57	3.072	3.3[illegible]7	0.	-.23[illegible]31700	-.371552537	.23722690	-.206[illegible]639	-.47	1.000
58	3.5[illegible]	3.392	0.	.18[illegible]95[illegible]00	.32597405	.13357360	-.7[illegible]131705	-1.53	1.000
59	3.[illegible]2	3.[illegible]	0.	-.1924[illegible]1400	-.3319[illegible]537	.23722640	-.1[illegible]0[illegible]6339	-.36	1.000
60	4.[illegible]	3.[illegible]32	0.	1.1[illegible]1[illegible]20	.3364[illegible]712	.14211710	.07520439	2.03	1.000
61	3.[illegible]3	3.71[illegible]	0.	.26[illegible]25300	1.3109555	.13[illegible]20[illegible]9	-.7[illegible]2[illegible]3930	-1.54	1.000
62	3.[illegible]	3.[illegible]2	0.	.[illegible]7137700	.73[illegible]5393	.14109014	-.32275536	-.67	1.000
63	3.[illegible]	2.[illegible]7	[illegible]	-1.03[illegible]6720	-1.22[illegible]5[illegible]7[illegible]	.[illegible]712[illegible]111	.17[illegible]6[illegible]32	1.03	1.000
64	3.[illegible]3	2.[illegible]	1.036	0.	-1.1042[illegible]13	.26513310	1.1542513	0.00	0.
65	4.171	2.23[illegible]	1.0[illegible]	0.	-2.[illegible]62412	.2629[illegible]469	2.4462512	0.00	0.
66	3.1[illegible]3	2.4[illegible]	1.000	0.	-2.3777[illegible]7	.27749574	2.0777637	0.00	0.
67	3.[illegible]1[illegible]	3.709	1.000	0.	-1.5523315	.41033632	1.6[illegible]23315	0.00	0.
68	4.[illegible]1[illegible]	3.[illegible]8	0.	0.	.3042331215	.24959324	-.3042331215	0.00	0.
69	4.[illegible]2	2.3[illegible]	1.000	0.	-.4766[illegible]0[illegible]	.27136406	.47660505	0.00	0.



THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

LEAST SQUARES FIT FOR DATA IN COLUMN 22
AS A LINEAR FUNCTION OF 21 PREDICTOR VARIABLES IN COLUMNS 1, 2, 3, 4, 5, 6, 7, 8,
9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21
USING 13 NON-ZERO WEIGHTS AND 6 ZERO WEIGHTS IN COLUMN 23

VARIANCE-COVARIANCE MATRIX OF THE ESTIMATED COEFFICIENTS

COLUMN	1	2	3	4	5	6	7
1	.25923092						
2	-.0357[illegible]3419	.0017558767					
3	.0147[illegible]919	-.0016[illegible]	.0119272200				
4	.054[illegible]15	-.02[illegible]	.017[illegible]	.37392935			
5	-.0350978[illegible]	-.0012[illegible]360	-.0069877131	.05256[illegible]17	.12330114		
6	-.219[illegible]389	.04[illegible]123	-.0374[illegible]936	-.51[illegible]691	-.0695006492	3.9203946	
7	.0417103590	4.35[illegible]32-04	.0013[illegible]79[illegible]	-.310[illegible]	-.02[illegible]585	.0325[illegible]971	.065909257
8	-.04[illegible]7[illegible]	-1.[illegible]-04	.0017[illegible]033	.00[illegible]275733	6.3526309-04	-1.3931056-05	-.0072701490
9	.0113[illegible]2[illegible]	-.0017[illegible]738	-.301[illegible]146527	.011334352	.0101[illegible]994	-.033010252	-.026316750
10	.01[illegible]	-.0275[illegible]213	.0317[illegible]266	.01[illegible]1233	.0211134744	.017[illegible]	-.0078[illegible]70470
11	-.013[illegible]01	.0075[illegible]	-.001[illegible]21923	-.013320575	-.0197330111	-.00[illegible]909	.0075[illegible]707
12	-.017[illegible]022-	.00[illegible]6[illegible]3	-.012[illegible]	-.0[illegible]1945	-.00[illegible]5548	.1005[illegible]12	1.1[illegible]-0[illegible]
13	.020[illegible]295	-.00[illegible]12[illegible]9	.013[illegible]4347	.03314552	.024[illegible]534	-.18557034	.0543[illegible]019
14	.0211566[illegible]	-.0079[illegible]	.003[illegible]2595	.15[illegible]1431	.01229[illegible]52	-.7[illegible]30[illegible]101	-.030912242
15	-.3233[illegible]	.0175[illegible]	.065[illegible]3714	-.30[illegible]933	-.110[illegible]	1.2915919	.071335131
16	-.03[illegible]	.0013[illegible]302	-.01[illegible]097	.05[illegible]4139	.017[illegible]19194	-.2233[illegible]	-.019[illegible]19890
17	-.00[illegible]	2.[illegible]-04	-.005[illegible]370312	.01[illegible]233	.0594512517	-.021917617	-.0[illegible]3275469[illegible]
18	.3[illegible]3792[illegible]	.00[illegible]214	-.0113[illegible]557	-.02[illegible]	.017[illegible]1630	.0393[illegible]913	-.02[illegible]2[illegible]42
19	-.03[illegible]	-.0017[illegible]2[illegible]0	.03[illegible]4207	.0[illegible]5[illegible]1	-.0130[illegible]542	-.023722633	.0158[illegible]1193
20	.01[illegible]	7.3[illegible]51-04	-.011[illegible]460	-.0355[illegible]214	.110922093	.0140[illegible]	-.0059516309
21	-.0[illegible]21173[illegible]	-.0011[illegible]201	.0073[illegible]1169[illegible]	.017[illegible]01[illegible]	-.0069[illegible]29339	-.0233[illegible]3002	.0374[illegible]973

COLUMN	8	9	10	11	12	13	14
8	.013956437						
9	-.0101[illegible]33795	.0321[illegible]04					
10	4.3[illegible]336-04	.004[illegible]541	.015937084				
11	-7.[illegible]-04	-.0[illegible]2[illegible]	-.0111[illegible]324	.021727333			
12	-3.[illegible]-0[illegible]	-2.[illegible]-0[illegible]	1.[illegible]3-04	3.[illegible]31252-04	.013[illegible]9647		
13	4.23[illegible]-0[illegible]	-.0031[illegible]	-.00[illegible]6743	.00[illegible]35[illegible]	-.0036[illegible]228	.07[illegible]09[illegible]4	
14	-.001[illegible]	.02[illegible]31	.01[illegible]513	-.015253451	-.025[illegible]1549	-.0554[illegible]1150	.41975163
15	.0[illegible]17	-.15[illegible]69	-.631[illegible]693	.031[illegible]	.0[illegible]2251	.011027157	-.57662793
16	-.0[illegible]17[illegible]035	.01[illegible]32	.0111[illegible]153	-.0[illegible]532	-.016[illegible]3097	-.04[illegible]302[illegible]5	.12795594
17	-1.6[illegible]2[illegible]-04	.157[illegible]	3.[illegible]-04	-2.[illegible]23-04	-.001[illegible]757	3.0351[illegible]-04	.0[illegible]0421969
18	-.03[illegible]	.01[illegible]63	.0072[illegible]7497	-.0[illegible]753[illegible]	.00[illegible]491952	-.5047[illegible]	.00[illegible]3530161
19	.3017[illegible]754	-.012[illegible]53	-.0[illegible]5932832	.0[illegible]3540215[illegible]	-.0019349316	.0[illegible]35121279	-.0037[illegible]718
20	-.051[illegible]30	.0[illegible]103[illegible]	.0[illegible]1[illegible]6063	-1.35[illegible]26-0[illegible]	5.1174136-04	-.[illegible]	.0131[illegible]34
21	7.3[illegible]516-04	-.00[illegible]739	-.0015[illegible]1252	.0011[illegible]65[illegible]	-.0121[illegible]94	.0[illegible]72	-.0324[illegible]570

COLUMN	15	16	17	18	19	20	21
15	1.4[illegible]932						
16	-.7[illegible]1[illegible]	.11[illegible]					
17	-.127-09927	.0[illegible]	.39[illegible]7665				
18	-.[illegible]173	.0[illegible]15	.00[illegible]30[illegible]	.0[illegible]30[illegible]56			
19	.0[illegible]37	-.[illegible]	-.101[illegible]7[illegible]	-.[illegible]78[illegible]	.916[illegible]3377		

20	-.040969959	.015133087	.0043691630	.017773443	-.0092225093	.025338642	
21	.0132j3291	-.011663549	-.0025077892	-.011414133	.0074795893	-.0170-5633	.013074732

ANALYSIS OF VARIANCE
-DEPENDENT ON ORDER VARIABLES ARE ENTERED, UNLESS VECTORS ARE ORTHOGONAL-

COLUMN	SS=RED. DUE TO COEF.	CUM. MS REDUCTION	D.F.	CUM. RESIDUAL MS	D.F.	F(COEF=0)	P(F)	F(COEFS=0)	P(F)
1	8.0226-46	8.9226-93	1	2.523-117	62	35.607	.000	29.249	0.000
2	36.127365	22.525087	2	1.372?2?2	61	143.361	.000	26.942	0.000
3	54.22756	33.9[illegible]51	3	1.10166?2	60	-215.158	0.000	22.920	0.000
4	3.3716594	27.151145	4	.33151455	59	37.159	.000	12.239	.000
5	15.137206	24.765750	5	.71620703	58	60.256	.000	10.771	.000
6	4.1832050	21.336-75	6	.55933057	57	10.520	.000	7.679	.000
7	2.6253666	18.577036	7	.55066151	56	6.179	.007	7.082	.000
8	2.5376711	15.572218	8	.52519.92	55	10.059	.003	7.014	.000
9	1.1[illegible]	14.151757	9	.54602351	54	4.546	.039	6.779	.000
10	1.07442?7	13.[illegible]33	10	.54860-6	53	4.264	.045	6.935	.000
11	.3556?-76	12.24[illegible]	11	.5[illegible]	52	1.461	.235	7.210	.000
12	.33[illegible]130	11.792787	12	.54062213	51	1.347	.253	7.786	.000
13	.0311[illegible]	10.[illegible]	13	.53731115	50	.007	.946	8.502	.000
14	1.8657[illegible]	9.412?541	14	.5?1[illegible]59	49	7.[illegible]	.009	9.564	.000
15	4.1495060	9.[illegible]5271	15	.43691155	48	13.450	.000	9.573	.000
16	.012[illegible]	9.3[illegible]	16	.50722 31	47	.048	.829	6.775	.000
17	.7[illegible]2?2	8.5710925	17	.50179 72	46	3.605	.190	10.520	.000
18	1.8032527	8.[illegible]319	18	.47274?00	45	7.130	.010	12.393	.000
19	.3938[illegible]7	7.5072253	19	.47[illegible]2?97	44	1.551	.216	14.139	.000
20	.6739[illegible]	7.24[illegible]	20	.4[illegible]79?12	43	2.573	.119	25.413	.000
21	9.6171-53	7.37[illegible]	21	.2?0008?0	42	34.153	.000	38.163	.000
RESIDUAL	16.534121		42						
TOTAL	165.37-18		63						

LEAST SQUARES FIT FOR DATA IN COLUMN 22
AS A LINEAR FUNCTION OF 21 PREDICTOR VARIABLES IN COLUMNS 1, 2, 3, 4, 5, 6, 7, 8,
9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21
USING 63 NON-ZERO WEIGHTS AND 6 ZERO WEIGHTS IN COLUMN 23

ESTIMATES FROM LEAST SQUARES FIT / FIT OMITTING LAST COLUMN

COLUMN	COEFFICIENT	S.D. OF COEFF.	RATIO	ACC. DIGITS	COEFFICIENT	S.D. OF COEFF.	RATIO
1	-.0315176	[illegible]	-8.04	12.80	-4.5352652	.6A622757	-6.59
2	[illegible]	[illegible]	3.01	13.02	.25612370	.11132517	2.29
3	[illegible]	[illegible]	2.02	12.65	[illegible]	.1003201	[illegible]
4	[illegible]	[illegible]	-7.79	13.03	-2.1295335	1.0255932	-2.08
5	[illegible]	[illegible]	-6.29	13.11	-2.5719339	.47219550	-5.45
6	[illegible]	[illegible]	2.21	12.87	3.1157736	2.1.883967	1.16
7	[illegible]	[illegible]	-1.45	12.66	.057035559	.3104909	.18
8	[illegible]	.11513735	-3.09	12.76	[illegible]	.16133564	-2.66
9	[illegible]	[illegible]	1.62	13.16	[illegible]	.22451676	1.52
10	[illegible]	[illegible]	2.01	13.11	[illegible]	.17032619	.90
11	[illegible]	[illegible]	-1.90	13.14	-.2304.7071	.20079128	-1.15
12	[illegible]	[illegible]	2.31	13.31	[illegible]	[illegible]	1.12
13	[illegible]	[illegible]	-3.42	13.05	[illegible]	[illegible]	-1.50
14	[illegible]	[illegible]	-1.70	12.75	-1.1032076	[illegible]	-1.25
15	[illegible]	[illegible]	2.22	13.04	3.7192619	[illegible]	2.29
16	[illegible]	[illegible]	1.51	13.49	[illegible]	[illegible]	-.05
17	[illegible]	[illegible]	3.15	13.30	.10400195	[illegible]	1.05
18	[illegible]	[illegible]	3.67	13.22	.089170716	[illegible]	.40
19	[illegible]	[illegible]	[illegible]	13.28	[illegible]	.13765454	-1.03
20	[illegible]	[illegible]	5.29	13.07	[illegible]	[illegible]	[illegible]
21	[illegible]	[illegible]	-6.19	13.02	-.091994336	.076927322	-1.20

RESIDUAL STANDARD DEVIATION = .50195535 .689.1748
BASED ON DEGREES OF FREEDOM 63-21 = 42 63-20 = 43

* THE NUMBER OF CORRECTLY COMPUTED DIGITS IN EACH COEFFICIENT USUALLY DIFFERS BY LESS THAN 1 FROM THE NUMBER GIVEN HERE
S AFTER FIT

LIST OF COMMANDS, DATA AND DIAGNOSTICS

[illegible]
[illegible]
[illegible] DIMENSION
[illegible] CARDS, IN COLS 2-[illegible]

[illegible]	[illegible]	0.000000	0.000000	0.000000	2.873565	6.010157	5.293317
10.[illegible]	3.61[illegible]	0.000000	0.000000	0.000000	3.600049	7.695244	5.16[illegible]
[illegible]	[illegible]	0.000000	0.000000	0.000000	3.660879	6.1954[illegible]	5.350566
10.[illegible]	3.[illegible]	0.000000	0.000000	0.000000	3.421000	4.3557[illegible]	5.709304
[illegible]	[illegible]	0.000000	0.000000	0.000000	3.658879	6.5[illegible]	4.347865
7.[illegible]	3.[illegible]712	0.000000	0.000000	0.000000	3.597312	6.630653	5.666427
[illegible]	1.33[illegible]	0.000000	0.000000	0.000000	1.350294	2.[illegible]7	2.991732
7.[illegible]1	.2231[illegible]	0.000000	0.000000	0.000000	.223144	[illegible]17	[illegible]
[illegible]	[illegible]	0.000000	0.000000	0.000000	2.639357	[illegible]	[illegible]
6.21[illegible]	[illegible]	0.000000	0.000000	1.000000	2.[illegible]1389	6.221923	5.5[illegible]643
[illegible]	1.671[illegible]	0.000000	0.000000	0.000000	1.671862	5.010635	2.991732
10.[illegible]	3.39[illegible]	0.000000	0.000000	[illegible]	3.39[illegible]508	3.[illegible]736	4.24[illegible]
7.[illegible]	2.214[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	4.01[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	4.40[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	.5[illegible]11	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	.137[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	2.7[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	4.77[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	[illegible]150[illegible]3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	1.2237[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	2.3[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8.[illegible]	4.04[illegible]75	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6.21[illegible]	[illegible]71534	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	[illegible]140[illegible]	[illegible]	0.000000	[illegible]	0.000000	0.000000	0.000000
12.[illegible]	.5[illegible]3	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
6.[illegible]	2.1[illegible]5	0.000000	0.000000	[illegible]	0.000000	0.000000	0.000000
10.[illegible]	3.4[illegible]775	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
12.11[illegible]	3.6[illegible]3	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
9.72[illegible]	3.[illegible]33[illegible]	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
7.[illegible]195	2.01[illegible]13	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.517953	4.04[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.779[illegible]	2.035[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]371	2.137[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	1.336[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
[illegible]	3.277[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9.[illegible]	[illegible]357[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]3	2.317[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]	2.73[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9.[illegible]12	3.218[illegible]	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000
12.3[illegible]	3.713772	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000
11.[illegible]	5.1[illegible]370	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
10.7[illegible]	3.[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12.[illegible]	[illegible]7[illegible]	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9.17[illegible]	3.5565[illegible]	0.000000	1.000000	1.000000	0.000000	0.000000	0.000000
[illegible]632	2.1[illegible]6	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000

5.734517	6.171761	0.000000	0.000000	0.000010	0.000000	0.000000	2.549445
10.[illegible]	7.174977	0.005000	19.459153	7.174647	3.455735	5.573730	3.459735
11.[illegible]	7.332327	0.000000	11.517415	7.720527	3.609373	6.745236	3.637373
[illegible]	4.059475	0.000000	0.000000	0.000000	0.000000	0.000000	3.433987
7.145256	5.553[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000	2.614903
7.4[illegible]	4.553145	0.000000	0.000000	0.000000	0.000000	0.000000	4.094345
7.373[illegible]	5.514311	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7.[illegible]371	5.513013	0.000000	0.000000	0.000000	0.000000	0.000000	2.197225
0.000000	0.000000	5.682131	0.000000	0.000000	0.000000	0.000000	1.363294
0.000000	0.000000	7.020427	0.000000	0.000000	0.000000	0.000000	3.777348
0.000000	0.000000	7.088951	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	3.[illegible]5349	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	5.714563	0.000000	0.000000	-0.000000	0.000000	2.713050
0.000000	0.000000	7.134579	9.355512	5.411775	3.216875	7.170120	0.000000
0.000000	0.000000	8.044531	12.382953	7.731043	3.713572	7.170120	0.000000
0.000000	0.000000	7.751155	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	[illegible]	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	2.456577	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	9.179354	7.473069	3.666562	5.657973	3.600062
0.000000	0.000000	2.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.277336
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	5.270916	0.000000	0.000000	0.000000	0.000000	2.075442
0.000000	0.000000	3.676492	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.583519
0.000000	0.000000	0.000000	0.364000	0.000000	0.000000	0.000000	3.583519
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.367296
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2.634057
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.591020
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.391020
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.367296
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.591020
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.367296
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.531926
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.713572
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.651820
0.000000	0.000000	0.000000	3.744743	5.459976	2.197225	3.555348	2.197225
0.000000	0.000000	3.774956	0.000000	0.000000	0.000000	0.000000	2.521734
0.000000	0.000000	4.650443	0.000000	0.000000	0.000000	0.000000	0.000000
0.000000	0.000000	5.742334	0.000000	0.000000	0.000000	0.000000	2.667000
0.000000	0.000000	4.292459	0.000000	0.000000	0.000105	0.000006	3.701652
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	3.768153
0.000000	0.000000	5.202355	0.000000	0.000000	0.000000	0.000000	2.944439

CR[illegible] TAPE C 69 CA 00, COLS 13[illegible]

2.673[illegible]	5.236317	0.000000	0.000000	-.521816
7.[illegible]	4.154[illegible]	3.600049	.229419	-.365723
[illegible]	0.000000	0.000000	0.000000	-.173775
7.[illegible]	6.7335[illegible]	0.000000	0.000000	1.302952
7.[illegible]	4.793815	0.000000	0.000000	.105272
7.[illegible]312	4.[illegible]27	0.000000	0.000000	-.900074

LIST OF COMMANDS, DATA AND DIAGNOSTICS

```
  [illegible]  4.495416    3.767295    4.573352   -.276317
  0.000000    0.000000    0.000000    0.000000    .185017
  [illegible]  4.449919    3.767296    4.573352   -.192414
  0.000000    0.000000    0.000000    0.000000    1.115112
  0.000000    0.000000    0.000000    0.000000    .2[illegible]8
  0.000000    0.000000    0.000000    0.000000    .471877
  0.000000    0.000000    0.000000    0.000000   -1.037672
  [illegible]  [illegible]  0.000000    0.000000    0.000000
  [illegible]  [illegible]  0.000000    0.000000    0.000000
  [illegible]  [illegible]  2.[illegible]009  [illegible]  0.000000
  [illegible]  [illegible]  3.[illegible]2  3.734552    0.000000
  [illegible]  [illegible]  0.000000    0.000000    0.000000
  [illegible]  [illegible]  0.000000    0.000000    0.000000
NOTE [illegible] AFTER READ
DEFINE 1.0 INTO COL 1
DEFINE 1.0 INTO COL 23
DEFINE 0.[illegible] INTO ROW [illegible] COL 23
DEFINE 0.9 INTO ROW [illegible] COL 23
DEFINE [illegible] INTO ROW [illegible] COL 23
DEFINE [illegible] INTO ROW [illegible] COL 23
DEFINE [illegible] INTO ROW 53 COL 23
DEFINE [illegible] INTO ROW 51 COL 23
FIT 22, [illegible] 23, 21 VARS, COL 1***21, RESID IN 24
NOTE . AFTER FIT
RESET [illegible] 43
PRINT COL 24
NOTE AFTER PRINT
STOP
```

NATIONAL BUREAU OF STANDARDS, WASH DC 20234, OMNITAB VERSION 5.00 MAY 15,1970
AIR FORCE INSTITUTE OF TECHNOLOGY, WPAFB OH, [illegible] VERSION 1.1 DEC 1, 1973

MIN FIELD LENGTH REQUIRED FOR THIS PROBLEM = [illegible]

Vita

Larry James Pulcher was born on September 18, 1951 in Belleville, Illinois. He graduated from high school in Dupo, Illinois, whereupon he attended Parks College of Aeronautical Technology of St. Louis University. He earned a degree there in Aerospace Engineering in 1972 and was commissioned into the Air Force through the ROTC program. His first assignment was as a Minuteman Launch Control Officer at Minot AFB, North Dakota where he was awarded a Regular commission and the Air Force Commendation Medal.

Permanent Address: R.R. 1, Box 265A
E. Carondelet, Illinois 62240

This thesis was typed by Mrs. Joyce Clark.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFIT/GOR/MA/78D-5
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): CRITERION FOR SELECTION OF VARIABLES IN A REGRESSION ANALYSIS
5. TYPE OF REPORT & PERIOD COVERED: M.S. Thesis
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Larry J. Pulcher, Captain USAF
8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Air Force Institute of Technology (AFIT/EN), Wright-Patterson AFB OH 45433
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Institute of Technology (AFIT/EN), Wright-Patterson AFB OH 45433
12. REPORT DATE: December 1978
13. NUMBER OF PAGES: 137
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES: Approved for public release; IAW AFR 190-17

JOSEPH P. HIPPS, Major, USAF
Director of Information
9 Jan 79

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Multiple Regression

Selection of Variables

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Test of Equality between Subsets of Coefficients in Two Regressions is developed and applied as a means to pre-screen variables from a regression model.

Some criterion for selection of variables are discussed and some existing regression packages are applied to data on characteristics of avionics equipment for comparison purposes.